

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2003, 09:22:57 ; Search time 14 Seconds  
(without alignments)  
487.580 Million cell updates/sec

Title: us-09-897-322-2

Perfect score: 1268

Sequence: 1 MERCPSLGVTLYALVVLVL.....QRGWIPIQYIISCKKSC 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5A COMB.pep:\*
- 2: /cgn2\_6/prodata/1/iaa/5B COMB.pep:\*
- 3: /cgn2\_6/prodata/1/iaa/6A COMB.pep:\*
- 4: /cgn2\_6/prodata/1/iaa/6B COMB.pep:\*
- 5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep:\*
- 6: /cgn2\_6/prodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1268	100.0	232	2	US-08-485-721-2
2	1268	100.0	232	2	US-08-392-935-2
3	1268	100.0	232	3	US-08-897-236-2
4	1268	100.0	232	4	US-09-167-874-2
5	1268	100.0	232	5	PCT-US93-08326-2
6	1258	99.2	232	3	US-08-897-236-11
7	1258	99.2	232	4	US-09-167-874-11
8	1195	94.2	449	3	US-08-897-236-23
9	949.5	74.9	222	2	US-08-485-721-9
10	949.5	74.9	222	2	US-08-392-935-9
11	949.5	74.9	222	5	PCT-US93-08326-9
12	503	39.7	87	2	US-08-485-721-11
13	503	39.7	87	2	US-08-392-935-11
14	503	39.7	87	5	PCT-US93-08326-11
15	93.5	7.4	372	3	US-08-586-165-5
16	92	7.3	538	4	US-09-309-572-12
17	92	7.3	1737	4	US-09-309-572-13
18	90.5	7.1	483	1	US-08-194-338-7
19	90.5	7.1	623	4	US-09-347-801-4
20	88.5	7.0	371	3	US-08-586-165-3
21	88.5	7.0	979	3	US-08-870-529-2
22	87	6.9	14	1	US-08-297-633A-3
23	87	6.9	14	2	US-08-485-721-3
24	87	6.9	14	2	US-08-392-935-3
25	87	6.9	14	3	US-08-897-236-3
26	87	6.9	14	4	US-09-167-874-3
27	87	6.9	14	5	PCT-US93-08325-3

28	87	6.9	14	5	PCT-US93-08326-3	Sequence 3, Appli
29	86	6.8	625	4	US-09-347-801-18	Sequence 18, Appl
30	85.5	6.7	315	4	US-08-957-351-27	Sequence 27, Appl
31	85	6.7	1621	1	US-08-242-677-2	Sequence 2, Appli
32	84	6.6	1593	4	US-08-628-829-4	Sequence 4, Appli
33	83	6.5	538	4	US-09-370-368-9	Sequence 9, Appli
34	79	6.2	11877	4	US-09-105-537-6	Sequence 6, Appli
35	76	6.0	14	1	US-08-297-633A-5	Sequence 5, Appli
36	76	6.0	14	2	US-08-485-721-5	Sequence 5, Appli
37	76	6.0	14	2	US-08-392-935-5	Sequence 5, Appli
38	76	6.0	14	3	US-08-897-236-5	Sequence 5, Appli
39	76	6.0	14	4	US-09-167-874-5	Sequence 5, Appli
40	76	6.0	14	5	PCT-US93-08325-5	Sequence 5, Appli
41	76	6.0	14	5	PCT-US93-08326-5	Sequence 5, Appli
42	75.5	6.0	6095	4	US-09-144-085-2	Sequence 2, Appli
43	74.5	5.9	438	3	US-08-486-099-105	Sequence 105, App
44	74.5	5.9	438	3	US-08-360-107A-115	Sequence 115, App
45	74.5	5.9	438	3	US-08-484-223B-105	Sequence 105, App

ALIGNMENTS

RESULT 1  
US-08-485-721-2  
; Sequence 2, Application US/08485721  
; Patent No. 5821124  
; GENERAL INFORMATION:  
; APPLICANT: Regeneron Pharmaceuticals, Inc. and  
; APPLICANT: Regents of the University of California  
; TITLE OF INVENTION: Dorsal Tissue Affecting Factor and  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
; STREET: 777 Old Saw Mill River Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,721  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/392,935  
; FILING DATE: 02-SEP-1993  
; APPLICATION NUMBER: PCT/US93/08326  
; FILING DATE: 02-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kempler Ph.D., Gail M.  
; REGISTRATION NUMBER: 32,143  
; REFERENCE/DOCKET NUMBER: Reg 132  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 914-347-7000  
; TELEFAX: 914-347-2113  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-485-721-2

Query Match 100.0%; Score 1268; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 2.3e-126;  
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MERCSLGTVLALVVLVGLRATPPGGGQYHLIRAPSDNIPLVYLIAHPPIIPKPKKD	60
Db	1	MERCSLGTVLALVVLVGLRATPPGGGQYHLIRAPSDNIPLVYLIAHPPIIPKPKKD	60
Qy	61	LNFTLLRSLTGGHYDPGFMAIISPPEDRRPGGGGAAGAEDLAELDQLRORPSGAMSEI	120
Db	61	LNFTLLRSLTGGHYDPGFMAIISPPEDRRPGGGGAAGAEDLAELDQLRORPSGAMSEI	120
Qy	121	KGLEFSEGLAGKKRRLSKTLRRKLIQMWLMSQTFCEPVLVYANNDGSRWPPIVYKGCSCS	180
Db	121	KGLEFSEGLAGKKRRLSKTLRRKLIQMWLMSQTFCEPVLVYANNDGSRWPPIVYKGCSCS	180
Qy	181	KRSGCVPEGMVCKPSKSVHLTVLMRCQRRGGORCGMIPLOYPIISSEKCGC	232
Db	181	KRSGCVPEGMVCKPSKSVHLTVLMRCQRRGGORCGMIPLOYPIISSEKCGC	232

RESULT 2  
US-08-392-935-2  
; Sequence 2, Application US/08392935

APPLICANT: Regeneron Pharmaceuticals, Inc. and  
 APPLICANT: Regents of the University of California  
 TITLE OF INVENTION: Dorsal Tissue Affecting Factor and  
 TITLE OF INVENTION: Compositions

ADDRESS: Regeneron Pharmaceuticals, Inc  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10591

```

? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Parentin Release #1.0, Version #1.2
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/392,935
? FILING DATE:

```

APPLICATION NUMBER: PCT/US93/083280  
FILING DATE: 02-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempler Ph.D., Gail M.  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: Reg 132  
TELECOMMUNICATION INFORMATION:

```

; INFORMATION FOR SEQ ID NO: 2
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 232 amino acids
;   TYPE: amino acid
;   TOPOLOGY: unknown
;
; MOLECULE TYPE: protein
;
;
;

```

```

Best Local Similarity 100.0%
Matches 232; Conservative

QY      1  MERCPSLGVTLYALVWVLGLRAA
        |||||||
Db       1  MERCPSLGVTLYALVWVLGLRAA

```

Db 61 LNETLRSLLGGHYDPGFMTSSE

Qy	121	KGLFEFEGSLAOGKORLSKRLRKLQWMLMSQIFCVLLVANDDLSRFPMPRYVKGSCFS	180
	121	KGLFEFEGSLAOGKORLSKRLRKLQWMLMSQIFCVLLVANDDLSRFPMPRYVKGSCFS	180
Db	121	KGLFEFEGSLAOGKORLSKRLRKLQWMLMSQIFCVLLVANDDLSRFPMPRYVKGSCFS	180
Qy	181	KRCSVPBGMVCKPSSVHLTVLWRCQRRGGRCCHPIQYPIISSECKSC	232
Db	181	KRCSVPBGMVCKPSSVHLTVLWRCQRRGGRCCHPIQYPIISSECKSC	232

RESULT 3  
US-08-897-236-2  
; Sequence 2, Application US/08897236A  
Data No. 007007

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 232
; TYPE: PRT
; ORGANISM: human
US-08-697-236-2

```

	Beet Local Similarity	100.0%; Pred. No. 2,3e-126;	Matches 233;	Conservative 0;	Mismatches 0;	Indels 0
Qy	1	MERCPSLGVTLVYALVWVIGLRATPAGQOHVLRIPAPSDNLPLVDLIERHDPPIFDE				
D6	1	MERCPSLGVTLVYALVWVIGLRATPAGQOHVLRIPAPSDNLPLVDLIERHDPPIFDE				

Qy	121	KGLEEBEGLAOGKQRILSKULRRKLQWMLWSQTFPCVLYAANDLGSFMPRYKVGSCFS	180
Db	61	LNNTLLRSLTGLGHYDPCGMATSPEDBPGGGGGAAGAEDLAEIQLRQRSGAMPEI	120
Qy	121	KGLEEBEGLAOGKQRILSKULRRKLQWMLWSQTFPCVLYAANDLGSFMPRYKVGSCFS	180
Db	61	LNNTLLRSLTGLGHYDPCGMATSPEDBPGGGGGAAGAEDLAEIQLRQRSGAMPEI	120

Db 181 KRCSVEGNCVCKPFSVHLTLVLRMCQRRGGQRCGMPIQIPIISECKSC 232

APPLICANT: VALENZUELA et al.  
TITLE OF INVENTION: DORSAL TISSUE AFFECTING FACTOR AND COMPOSITIONS  
FILE REFERENCE: REG332-B  
CURRENT APPLICATION NUMBER: US/09/167,874  
CURRENT FILING DATE: 1998-10-07  
EARLIER APPLICATION NUMBER: 08/485,721  
PUBLICATION NUMBER: 08/485,721

```

1 EARLIER APPLICATION NUMBER PCT/US93/0832
2 EARLIER FILING DATE: 1993-09-02
3 EARLIER APPLICATION NUMBER: 07/957,401
4 EARLIER FILING DATE: 1992-10-06
5 EARLIER APPLICATION NUMBER: 07/950,410
6 EARLIER FILING DATE: 1992-09-23

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

```

```

; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-874-2

Query Match          100.0%; Score 1268; DB 4; Length 232;
Best Local Similarity 100.0%; Pred. No. 2.3e-126;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYHLIRPAPSDNLPLVDLIEHDPDPIFDPKEKD 60
Db 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYHLIRPAPSDNLPLVDLIEHDPDPIFDPKEKD 60

QY 61 LNETLLRSLGGHYDPGFMATSPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120
Db 61 LNETLLRSLGGHYDPGFMATSPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

QY 121 KGLEFSEGLAOGKKORLSKKLRKQLMWLWSOTFCPVLYAMNDLGSRFWPRYVKVGCSCFS 180
Db 121 KGLEFSEGLAOGKKORLSKKLRKQLMWLWSOTFCPVLYAMNDLGSRFWPRYVKVGCSCFS 180

QY 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGQRCGWIPQIYPIIIECKKSC 232
Db 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGQRCGWIPQIYPIIIECKKSC 232

RESULT 5
PCT-US93-08326-2
; Sequence 2, Application PC/TUS9308326
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: Dorsal Tissue Affecting Factor and
; TITLE OF INVENTION: Compositions
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08326
; FILING DATE: 02-SEP-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler Ph.D., Gail M.
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: Reg 132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-347-7000
; TELEFAX: 914-347-2113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US93-08326-2

Query Match          100.0%; Score 1268; DB 5; Length 232;
Best Local Similarity 100.0%; Pred. No. 2.3e-126;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYHLIRPAPSDNLPLVDLIEHDPDPIFDPKEKD 60
Db 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYHLIRPAPSDNLPLVDLIEHDPDPIFDPKEKD 60
```

```

QY 61 LNETLLRSLGGHYDPGFMATSPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120
Db 61 LNETLLRSLGGHYDPGFMATSPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

QY 121 KGLEFSEGLAOGKKORLSKKLRKQLMWLWSOTFCPVLYAMNDLGSRFWPRYVKVGCSCFS 180
Db 121 KGLEFSEGLAOGKKORLSKKLRKQLMWLWSOTFCPVLYAMNDLGSRFWPRYVKVGCSCFS 180

QY 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGQRCGWIPQIYPIIIECKKSC 232
Db 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGQRCGWIPQIYPIIIECKKSC 232

RESULT 6
US-08-897-236-11
; Sequence 11, Application US/08897236A
; Patent No. 6075007
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modified Dorsal Tissue Affecting Factor and Composition
; FILE REFERENCE: REG 133
; CURRENT APPLICATION NUMBER: US/08/897,236A
; CURRENT FILING DATE: 1997-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 232
; TYPE: PRT
; ORGANISM: mouse
US-08-897-236-11

Query Match          99.2%; Score 1258; DB 3; Length 232;
Best Local Similarity 99.1%; Pred. No. 2.7e-125;
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYHLIRPAPSDNLPLVDLIEHDPDPIFDPKEKD 60
Db 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYHLIRPAPSDNLPLVDLIEHDPDPIFDPKEKD 60

QY 61 LNETLLRSLGGHYDPGFMATSPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120
Db 61 LNETLLRSLGGHYDPGFMATSPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

QY 121 KGLEFSEGLAOGKKORLSKKLRKQLMWLWSOTFCPVLYAMNDLGSRFWPRYVKVGCSCFS 180
Db 121 KGLEFSEGLAOGKKORLSKKLRKQLMWLWSOTFCPVLYAMNDLGSRFWPRYVKVGCSCFS 180

QY 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGQRCGWIPQIYPIIIECKKSC 232
Db 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCQRRGQRCGWIPQIYPIIIECKKSC 232

RESULT 7
US-09-167-874-11
; Sequence 11, Application US/09167874
; Patent No. 6277593
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: DORSAL TISSUE AFFECTING FACTOR AND COMPOSITIONS
; FILE REFERENCE: REG132-B
; CURRENT APPLICATION NUMBER: US/09/167,874
; CURRENT FILING DATE: 1998-10-07
; EARLIER APPLICATION NUMBER: 08/485,721
; EARLIER FILING DATE: 1995-07-06
; EARLIER APPLICATION NUMBER: 08/392,935
; EARLIER FILING DATE: 1995-09-22
; EARLIER APPLICATION NUMBER: PCT/US93/08326
; EARLIER FILING DATE: 1993-09-02
; EARLIER APPLICATION NUMBER: 07/957,401
; EARLIER FILING DATE: 1992-10-06
; EARLIER APPLICATION NUMBER: 07/950,410
; EARLIER FILING DATE: 1992-09-23
; EARLIER APPLICATION NUMBER: 07/939,954
```



```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392,935
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08326
; FILING DATE: 02-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler Ph.D., Gail M.
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: Reg 132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-347-7000
; TELEFAX: 914-347-2113
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-392-935-9

Query Match 74.9%; Score 949.5; DB 2; Length 222;
Best Local Similarity 78.1%; Pred. No. 1.2e-92;
Matches 175; Conservative 17; Mismatches 23; Indels 9; Gaps 2;

QY 9 VTIYALVVVLGLRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDPIDPKKOLNETLLRS 68
Db 8 VTIIYALMVFLGLRIDGGQHYLHIRPAPSENLPVLDLIEHPDPIDPKKOLNETLLRT 67
QY 69 ILGGHYDPCFMATSPEDRPGGGGGAAGAEADLAELDQLLRQPSGAMPSEIKGLEFSEG 128
Db 68 LMVGHDPNFMATILPEERL-----GVEDLGELDLLLRQKPSGAMPAPAEIKGLEFYEG 119
QY 129 LAQKKORLSKRLRRKLQWLWSQTFCPVLYAWNDLGSFRWPRYVKVGSFCFSKRSVPE 188
Db 120 L-QSKHRLSKLRRKLQWLWSQTFCPVLYTWNLDGTRFWPRYVKVGSYKRSVPE 178
QY 189 GMVCKPSKSVHLTVLRWCRRGQCGWIPQIYPIISECKSC 232
Db 179 GNVCKAAKSMHILRWRCRRVQKCAWITIYQPVISECKSC 222

RESULT 11
PCT-US93-08326-9
; Sequence 9, Application PC/TUS9308326
; GENERAL INFORMATION:
; APPLICANT: Valenzuela, et al.
; TITLE OF INVENTION: Dorsal Tissue Affecting Factor and
; TITLE OF INVENTION: Compositions
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08326
; FILING DATE: 02-SEP-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kempler Ph.D., Gail M.
; REGISTRATION NUMBER: 32,143
; REFERENCE/DOCKET NUMBER: Reg 132
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 914-347-7000
; TELEFAX: 914-347-2113
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US93-08326-9

Query Match 74.9%; Score 949.5; DB 5; Length 222;
Best Local Similarity 78.1%; Pred. No. 1.2e-92;
Matches 175; Conservative 17; Mismatches 23; Indels 9; Gaps 2;

QY 9 VTIYALVVVLGLRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDPIDPKKOLNETLLRS 68
Db 8 VTIIYALMVFLGLRIDGGQHYLHIRPAPSENLPVLDLIEHPDPIDPKKOLNETLLRT 67
QY 69 ILGGHYDPCFMATSPEDRPGGGGGAAGAEADLAELDQLLRQPSGAMPSEIKGLEFSEG 128
Db 68 LMVGHDPNFMATILPEERL-----GVEDLGELDLLLRQKPSGAMPAPAEIKGLEFYEG 119
QY 129 LAQKKORLSKRLRRKLQWLWSQTFCPVLYAWNDLGSFRWPRYVKVGSFCFSKRSVPE 188
Db 120 L-QSKHRLSKLRRKLQWLWSQTFCPVLYTWNLDGTRFWPRYVKVGSYKRSVPE 178
QY 189 GMVCKPSKSVHLTVLRWCRRGQCGWIPQIYPIISECKSC 232
Db 179 GNVCKAAKSMHILRWRCRRVQKCAWITIYQPVISECKSC 222

RESULT 12
US-08-485-721-11
; Sequence 11, Application US/08485721
; Patent No. 5821124
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc. and
; APPLICANT: Regents of the University of California
; TITLE OF INVENTION: Dorsal Tissue Affecting Factor and
; TITLE OF INVENTION: Compositions
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.
; STREET: 777 Old Saw Mill River Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,721
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/392,935
; FILING DATE: 02-SEP-1993
; APPLICATION NUMBER: PCT/US93/08326
; FILING DATE: 02-SEP-1993
```

ATTORNEY/AGENT INFORMATION:  
NAME: Kempster Ph.D., Gail M.  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: Reg 132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-347-7000  
TELEFAX: 914-347-2113  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-485-721-11

Query Match 39.7%; Score 503; DB 2; Length 87;  
Best Local Similarity 98.9%; Pred. No. 5.9e-46;  
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 146 QMWLMSQTFPCPVLYAMNDLSRFPWRYVKGSCFSKRSQSVBEGWCKPSKSVLTVLRW 205  
Db 1 QMWLMSQTFPCPVLYAMNDLSRFPWRYVKGSCFSKRSQSVBEGWCKPSKSVLTVLRW 60  
Qy 206 RCORRGORCGWIPIOYPIISECKCSC 232  
Db 61 RCORRGORCGWIPIOYPIISECKCSC 87

RESULT 13  
US-08-392-935-11  
Sequence 11, Application US/08392935  
Patent No. 5843775  
GENERAL INFORMATION:  
APPLICANT: Regeneron Pharmaceuticals, Inc. and  
APPLICANT: Regents of the University of California  
TITLE OF INVENTION: Dorsal Tissue Affecting Factor and  
TITLE OF INVENTION: Compositions  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,935  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08326  
FILING DATE: 02-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempster Ph.D., Gail M.  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: Reg 132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-347-7000  
TELEFAX: 914-347-2113  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-392-935-11

Query Match 39.7%; Score 503; DB 2; Length 87;

Best Local Similarity 98.9%; Pred. No. 5.9e-46;  
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 146 QMWLMSQTFPCPVLYAMNDLSRFPWRYVKGSCFSKRSQSVBEGWCKPSKSVLTVLRW 205  
Db 1 QMWLMSQTFPCPVLYAMNDLSRFPWRYVKGSCFSKRSQSVBEGWCKPSKSVLTVLRW 60  
Qy 206 RCORRGORCGWIPIOYPIISECKCSC 232  
Db 61 RCORRGORCGWIPIOYPIISECKCSC 87

RESULT 14  
PCT-US93-08326-11  
Sequence 11, Application PC/TUS9308326  
GENERAL INFORMATION:  
APPLICANT: Valenzuela, et al.  
TITLE OF INVENTION: Dorsal Tissue Affecting Factor and  
TITLE OF INVENTION: Compositions  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/08326  
FILING DATE: 02-SEP-1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempster Ph.D., Gail M.  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: Reg 132  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-347-7000  
TELEFAX: 914-347-2113  
INFORMATION FOR SEQ. ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US93-08326-11

Query Match 39.7%; Score 503; DB 5; Length 87;  
Best Local Similarity 98.9%; Pred. No. 5.9e-46;  
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 146 QMWLMSQTFPCPVLYAMNDLSRFPWRYVKGSCFSKRSQSVBEGWCKPSKSVLTVLRW 205  
Db 1 QMWLMSQTFPCPVLYAMNDLSRFPWRYVKGSCFSKRSQSVBEGWCKPSKSVLTVLRW 60  
Qy 206 RCORRGORCGWIPIOYPIISECKCSC 232  
Db 61 RCORRGORCGWIPIOYPIISECKCSC 87

RESULT 15  
US-08-586-165-5  
Sequence 5, Application US/08586165  
Patent No. 6054298  
GENERAL INFORMATION:  
APPLICANT: Laufer, Edward M.  
APPLICANT: Orozco, Olivia E.  
APPLICANT: Tabin, Clifford J.  
TITLE OF INVENTION: Fringe Proteins and Pattern Formation

NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/586,165  
FILING DATE: 16-JAN-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: H095-05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-586-165-5

Query Match 7.4%; Score 93.5; DB 3; Length 372;  
Best Local Similarity 25.4%; Pred No. 0.1;  
Matches 52; Conservative 17; Mismatches 81; Indels 55; Gaps 11;  
Qy 4 CPSLGVTLVYLVGLRATPAGGQHYLHIRPAPSDNLELVDLIEHPDIFD--PKEKDL 61  
Db 12 CPLLSTAAVLELLLPQGPPAAPP--RPPAGP-----SRPSKREARPAGSDV 61  
Qy 62 NETLLRSLGCHYDFGMATSPDEPFGGGGAAGGAEDLAE---DQLLRORPSGMP 118  
Db 62 -----PGDR-GGSGAAGGGRVAGSPPSRRVRMGPPGGSAAK 98  
Qy 119 EIKGLEFSEGLAQKKORLSKKLRKQMWLW-----SOTFCPLVYANDLGSRFWRPV 173  
Db 99 E--SLELKDIFIAVTKTKYKHTRLELLFQTWISRAGQT--IFTDWEDELRL-----L 149  
Qy 174 KVGSCFSKRSCS---VPEGMVCKPS 195  
Db 150 KAGDHMTNCSAVHTQALCKMS 174

Search completed: January 7, 2003, 09:25:20  
Job time : 15 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2003, 09:24:43 ; Search time 10 Seconds  
(without alignments)  
439.675 Million cell updates/sec

Title: US-09-897-322-2  
Perfect score: 1268  
Sequence: 1 MERCPSLGVTLYALVVVLGL.....QRCGWIPQYPIISECKKSC 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 117078 segs, 18951520 residues  
Total number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/prodata/1/pubpaa/US08\_NEW PUB.pep.\*
- 2: /cgn2\_6/prodata/1/pubpaa/FCI\_NEW PUB.pep.\*
- 3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW PUB.pep.\*
- 4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW PUB.pep.\*
- 6: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/1/pubpaa/FCIUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/1/pubpaa/US09\_NEW PUB.pep.\*
- 10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/1/pubpaa/US10\_NEW PUB.pep.\*
- 12: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/prodata/1/pubpaa/US60\_NEW PUB.pep.\*
- 14: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1268	100.0	232	9	US-10-044-716-4
2	1268	100.0	232	10	US-09-897-322-2
3	1258	99.2	232	9	US-10-044-716-6
4	1258	99.2	232	10	US-09-897-322-11
5	90.5	7.1	623	10	US-09-854-731-4
6	88.5	7.0	979	9	US-09-884-001-4
7	87	6.9	14	10	US-09-897-322-3
8	86.5	6.8	939	10	US-09-854-845-16
9	86.5	6.8	954	10	US-09-854-845-14
10	86.5	6.8	1034	10	US-09-854-845-6
11	86.5	6.8	1049	10	US-09-854-845-2
12	86.5	6.8	1078	10	US-09-854-845-8
13	86.5	6.8	1093	10	US-09-854-845-4
14	86.5	6.8	1136	10	US-09-854-845-12
15	86.5	6.8	1151	10	US-09-854-845-10
16	86	6.8	625	10	US-09-854-731-18
17	83	6.5	538	10	US-09-827-822-9
18	80.5	6.3	465	9	US-10-029-180-44
19	80.5	6.3	938	9	US-10-029-217A-28

20	80.5	6.3	938	9	US-10-029-217A-29	Sequence 29, Appl
21	79	6.2	264	9	US-09-738-626-5623	Sequence 5623, Ap
22	79	6.2	325	9	US-10-063-547-64	Sequence 64, Appl
23	79	6.2	325	10	US-09-741-711-2	Sequence 2, Appl
24	79	6.2	325	12	US-10-006-867-64	Sequence 64, Appl
25	79	6.2	325	12	US-10-052-586-238	Sequence 238, App
26	79	6.2	11877	9	US-09-860-846-6	Sequence 6, Appl
27	79	6.2	11877	10	US-09-861-289-6	Sequence 6, Appl
28	78	6.2	135	10	US-09-925-297-852	Sequence 852, App
29	77	6.1	428	9	US-09-734-329-2	Sequence 2, Appl
30	77	6.1	1249	9	US-09-964-899-33	Sequence 33, Appl
31	76.5	6.0	144	10	US-09-811-284-131	Sequence 131, App
32	76	6.0	14	10	US-09-897-322-5	Sequence 5, Appl
33	76	6.0	169	10	US-09-823-038A-53	Sequence 53, Appl
34	76	6.0	588	10	US-09-782-980-78	Sequence 78, Appl
35	74.5	5.9	321	10	US-09-925-301-938	Sequence 938, App
36	74.5	5.9	550	10	US-09-873-233A-20	Sequence 20, Appl
37	74	5.8	848	9	US-09-882-986-2	Sequence 2, Appl
38	74	5.8	4613	9	US-09-860-846-31	Sequence 31, Appl
39	74	5.8	4613	10	US-09-861-289-31	Sequence 31, Appl
40	73.5	5.8	127	10	US-09-795-006A-47	Sequence 47, Appl
41	73.5	5.8	295	9	US-09-978-295A-2	Sequence 2, Appl
42	73.5	5.8	295	9	US-09-978-697-2	Sequence 2, Appl
43	73.5	5.8	295	9	US-09-978-132A-2	Sequence 2, Appl
44	73.5	5.8	295	9	US-09-999-832A-2	Sequence 2, Appl
45	73	5.8	12	10	US-09-897-322-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-10-044-716-4  
; Sequence 4, Application US/10044716  
; Patent No. US20020159986A1  
; GENERAL INFORMATION:  
; APPLICANT: LANGENFELD, John  
; FILE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF CA  
; TITLE REFERENCE: 270/070US  
; CURRENT APPLICATION NUMBER: US/10/044,716  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: US60/261,252  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 232  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-044-716-4

Query Match	100.0%	Score 1268;	DB 9;	Length 232;
Best Local Similarity	100.0%;	Pred. No. 2.8e-108;	Mismatches 0;	Indels 0;
Matches 232;	Conservative	0;	0;	Gaps 0;
Qy	1	MERCPSLGVTLYALVVVLGLRATPAGGQHYLHIRPAPSDNPLVDLIEHPDIFDPKED	60	
Db	1	MERCPSLGVTLYALVVVLGLRATPAGGQHYLHIRPAPSDNPLVDLIEHPDIFDPKED	60	
Qy	61	LNETLRSLLGGHYDPGFWMATSPEDRPGCGGGAAGGAEDLAELDOLLQRPSGAMPSEI	120	
Db	61	LNETLRSLLGGHYDPGFWMATSPEDRPGCGGGAAGGAEDLAELDOLLQRPSGAMPSEI	120	
Qy	121	KGLESEGLAOCKKORLSKKLRKLMWLSOTFCPLVYAWNDLGSRRFWRPVYKVGSCFS	180	
Db	121	KGLESEGLAOCKKORLSKKLRKLMWLSOTFCPLVYAWNDLGSRRFWRPVYKVGSCFS	180	
Qy	181	KRSCSVPEGMVCKPSKSVHLTVLRWRCQRGGRCQRCWIPQYPIISECKKSC 232		
Db	181	KRSCSVPEGMVCKPSKSVHLTVLRWRCQRGGRCQRCWIPQYPIISECKKSC 232		
RESULT 2				

```
US-09-897-322-2
; Sequence 2, Application US/09897322
; Patent No. US20020102643A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: DORSAL TISSUE AFFECTING FACTOR AND COMPOSITIONS
; FILE REFERENCE: REG132-B
; CURRENT APPLICATION NUMBER: US/09/897,322
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/167,874
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 08/485,721
; PRIOR FILING DATE: 1995-07-06
; PRIOR APPLICATION NUMBER: 08/392,935
; PRIOR FILING DATE: 1995-09-22
; PRIOR APPLICATION NUMBER: PCT/US93/08326
; PRIOR FILING DATE: 1993-09-02
; PRIOR APPLICATION NUMBER: 07/957,401
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: 07/950,410
; PRIOR FILING DATE: 1992-09-23
; PRIOR APPLICATION NUMBER: 07/939,954
; PRIOR FILING DATE: 1992-09-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-322-2
```

```
Query Match 100.0%; Score 1268; DB 10; Length 232;
Best Local Similarity 100.0%; Pred. No. 2,8e-108;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MERCPISLGVTLVYLVVVLGIRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDPKED 60
DB 1 MERCPISLGVTLVYLVVVLGIRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDPKED 60

QY 61 LNETLLRSLLGGHYDPGFMAATSPEDRPGGGGAAGAEADLAEIDQLLRQPSGAMPSEI 120
DB 61 LNETLLRSLLGGHYDPGFMAATSPEDRPGGGGAAGAEADLAEIDQLLRQPSGAMPSEI 120

QY 121 KGLEFSEGLAOGKKORLSKLRKLQMWLMSQTFPCVLYAMNDLSGRFWRVYKGCSCS 180
DB 121 KGLEFSEGLAOGKKORLSKLRKLQMWLMSQTFPCVLYAMNDLSGRFWRVYKGCSCS 180

QY 181 KRCSVPEGWVCKPSKSVHLTVLRMRCQRRGRCGMIPDIQYPIISECKCSC 232
DB 181 KRCSVPEGWVCKPSKSVHLTVLRMRCQRRGRCGMIPDIQYPIISECKCSC 232
```

```
RESULT 3
US-10-044-716-6
; Sequence 6, Application US/10044716
; Patent No. US2002015986A1
; GENERAL INFORMATION:
; APPLICANT: LANGENFELD, John
; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF CA
; FILE REFERENCE: 270/070US
; CURRENT APPLICATION NUMBER: US/10/044,716
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US60/261,252
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-044-716-6
```

Query Match 99.2%; Score 1258; DB 9; Length 232;

```
Best Local Similarity 99.1%; Pred. No. 2.3e-107;
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERCPISLGVTLVYLVVVLGIRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDPKED 60
DB 1 MERCPISLGVTLVYLVVVLGIRAPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDPKED 60

QY 61 LNETLLRSLLGGHYDPGFMAATSPEDRPGGGGAAGAEADLAEIDQLLRQPSGAMPSEI 120
DB 61 LNETLLRSLLGGHYDPGFMAATSPEDRPGGGGAGAEADLAEIDQLLRQPSGAMPSEI 120

QY 121 KGLEFSEGLAOGKKORLSKLRKLQMWLMSQTFPCVLYAMNDLSGRFWRVYKGCSCS 180
DB 121 KGLEFSEGLAOGKKORLSKLRKLQMWLMSQTFPCVLYAMNDLSGRFWRVYKGCSCS 180

QY 181 KRCSVPEGWVCKPSKSVHLTVLRMRCQRRGRCGMIPDIQYPIISECKCSC 232
DB 181 KRCSVPEGWVCKPSKSVHLTVLRMRCQRRGRCGMIPDIQYPIISECKCSC 232
```

```
RESULT 4
US-09-897-322-11
; Sequence 11, Application US/09897322
; Patent No. US20020102643A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: DORSAL TISSUE AFFECTING FACTOR AND COMPOSITIONS
; FILE REFERENCE: REG132-B
; CURRENT APPLICATION NUMBER: US/09/897,322
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/167,874
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 08/485,721
; PRIOR FILING DATE: 1995-07-06
; PRIOR APPLICATION NUMBER: 08/392,935
; PRIOR FILING DATE: 1995-09-22
; PRIOR APPLICATION NUMBER: PCT/US93/08326
; PRIOR FILING DATE: 1993-09-02
; PRIOR APPLICATION NUMBER: 07/957,401
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: 07/950,410
; PRIOR FILING DATE: 1992-09-23
; PRIOR APPLICATION NUMBER: 07/939,954
; PRIOR FILING DATE: 1992-09-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 232
; TYPE: PRT
; ORGANISM: mouse
US-09-897-322-11
```

```
Query Match 99.2%; Score 1258; DB 10; Length 232;
Best Local Similarity 99.1%; Pred. No. 2.3e-107;
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERCPISLGVTLVYLVVVLGIRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDPKED 60
DB 1 MERCPISLGVTLVYLVVVLGIRAPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDPKED 60

QY 61 LNETLLRSLLGGHYDPGFMAATSPEDRPGGGGAAGAEADLAEIDQLLRQPSGAMPSEI 120
DB 61 LNETLLRSLLGGHYDPGFMAATSPEDRPGGGGAGAEADLAEIDQLLRQPSGAMPSEI 120

QY 121 KGLEFSEGLAOGKKORLSKLRKLQMWLMSQTFPCVLYAMNDLSGRFWRVYKGCSCS 180
DB 121 KGLEFSEGLAOGKKORLSKLRKLQMWLMSQTFPCVLYAMNDLSGRFWRVYKGCSCS 180

QY 181 KRCSVPEGWVCKPSKSVHLTVLRMRCQRRGRCGMIPDIQYPIISECKCSC 232
DB 181 KRCSVPEGWVCKPSKSVHLTVLRMRCQRRGRCGMIPDIQYPIISECKCSC 232
```

Query Match 99.2%; Score 1258; DB 9; Length 232;

[illegible]

```

Db      685  TPTPRPHGSD-----ICTSWPRIFGSLHHVPDLSRCRWHTTILIVKVLNSKTIIRNSSG 733
Qy      73   -HYDPGFWATSPPEDRPGGGGGAAGAAEDLAEIDLQLLRORPSGAMPSEIKGLFSEGL-- 129
Db      740  LSICTVTFQSSS-----PGGGGGGGGEEEDSQESETPD-PSGGFRGTMEADRGMEGLIS 793
Qy      130  ---AOGKKQRLSKLKRKLQMWLWSQTFPCVLYAWNDLSGRF 168
Db      794  PTEAMGNSNGASSCPGWLKRKELENAEFIMPDPSPPLSAAF 835

RESULT 7
US-09-897-322-3
; Sequence 3, Application US/09897322
; Patent No. US20020102643A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al.
; TITLE OF INVENTION: DORSAL TISSUE AFFECTING FACTOR AND COMPOSITIONS
; FILE REFERENCE: REG132-B
; CURRENT APPLICATION NUMBER: US/09/897,322
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/167,874
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 08/485,721
; PRIOR FILING DATE: 1995-07-06
; PRIOR APPLICATION NUMBER: 08/392,935
; PRIOR FILING DATE: 1995-09-22
; PRIOR APPLICATION NUMBER: PCT/US93/08326
; PRIOR FILING DATE: 1993-09-02
; PRIOR APPLICATION NUMBER: 07/957,401
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: 07/950,410
; PRIOR FILING DATE: 1992-09-23
; PRIOR APPLICATION NUMBER: 07/939,954
; PRIOR FILING DATE: 1992-09-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 14
; TYPE: PRT
; ORGANISM: frog and mouse
US-09-897-322-3

Query Match      6.9%; Score 87; DB 10; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      146  QMWLWSQTFPCVLY 159
Db      1    QMWLWSQTFPCVLY 14

RESULT 8
US-09-854-845-16
; Sequence 16, Application US/09854845
; Patent No. US20020098491A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020098491A1a1 Human Semaphorin Homologs and
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 939

```

```
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-16

Query Match
Best Local Similarity 30.2%; Pred. No. 5.5;
Matches 35; Conservative 10; Mismatches 44; Indels 27; Gaps 9;

Oy 132 GKKORLSKLRKRLQWMLWSQ--TFCPVLYAMNDLGSRFMPRYV-----KVGSCFSK 181
Db 380 GKQORCS-TLEDSSNMSLWNTQNTACPVNRVTRDGGFGPWSWPQCEHLDGNSGCLCR 438

Oy 182 -RSCSVPE---GMVCKPSKSVHLTVLRMRCQRRGQRCGMPIQYPIISECKCSC 232
Db 439 ARSCDSPPRCGJLDCL-GPAIHIA---NCSRNG---AWTP--WSSWALCSTSC 483

RESULT 9
US-09-854-845-14
; Sequence 14, Application US/09854845
; Patent No. US20020098491A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotid
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 954
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-14

Query Match
Best Local Similarity 30.2%; Pred. No. 5.6;
Matches 35; Conservative 10; Mismatches 44; Indels 27; Gaps 9;

Oy 132 GKKORLSKLRKRLQWMLWSQ--TFCPVLYAMNDLGSRFMPRYV-----KVGSCFSK 181
Db 380 GKQORCS-TLEDSSNMSLWNTQNTACPVNRVTRDGGFGPWSWPQCEHLDGNSGCLCR 438

Oy 182 -RSCSVPE---GMVCKPSKSVHLTVLRMRCQRRGQRCGMPIQYPIISECKCSC 232
Db 439 ARSCDSPPRCGJLDCL-GPAIHIA---NCSRNG---AWTP--WSSWALCSTSC 483

RESULT 10
US-09-854-845-6
; Sequence 6, Application US/09854845
; Patent No. US20020098491A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotid
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-6

Query Match
Best Local Similarity 30.2%; Pred. No. 6.2;
Matches 35; Conservative 10; Mismatches 44; Indels 27; Gaps 9;

Oy 132 GKKORLSKLRKRLQWMLWSQ--TFCPVLYAMNDLGSRFMPRYV-----KVGSCFSK 181
Db 475 GKQORCS-TLEDSSNMSLWNTQNTACPVNRVTRDGGFGPWSWPQCEHLDGNSGCLCR 533

Oy 182 -RSCSVPE---GMVCKPSKSVHLTVLRMRCQRRGQRCGMPIQYPIISECKCSC 232
Db 534 ARSCDSPPRCGJLDCL-GPAIHIA---NCSRNG---AWTP--WSSWALCSTSC 578

RESULT 11
US-09-854-845-2
; Sequence 2, Application US/09854845
; Patent No. US20020098491A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotid
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-2

Query Match
Best Local Similarity 30.2%; Pred. No. 6.3;
Matches 35; Conservative 10; Mismatches 44; Indels 27; Gaps 9;

Oy 132 GKKORLSKLRKRLQWMLWSQ--TFCPVLYAMNDLGSRFMPRYV-----KVGSCFSK 181
Db 475 GKQORCS-TLEDSSNMSLWNTQNTACPVNRVTRDGGFGPWSWPQCEHLDGNSGCLCR 533

Oy 182 -RSCSVPE---GMVCKPSKSVHLTVLRMRCQRRGQRCGMPIQYPIISECKCSC 232
Db 534 ARSCDSPPRCGJLDCL-GPAIHIA---NCSRNG---AWTP--WSSWALCSTSC 578

RESULT 12
US-09-854-845-8
; Sequence 8, Application US/09854845
; Patent No. US20020098491A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotid
; FILE REFERENCE: LEX-0177-USA
; CURRENT APPLICATION NUMBER: US/09/854,845
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
```

```

; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-854-845-8

```

Query Match	6.8%	Score 86.5;	DB 10;	Length 1078;
Best Local Similarity	30.2%	Pred. No. 6.5;		
Matches 35;	Conservative	10;	Mismatches 44;	Indels 27;
				Gaps 9;

Qy	132	GKQRLSKLRRKLQWLWNSQ---TFQPVLYAWNDLGRFWRPVV-----KVGSCFSK 181
Db	519	GKQORCS-TLEDSSNLSLWNTITATCPVRNVTVDGGFGPSPQPCHELDGNDGNSCLCR 577
Qy	182	-RSCSYPE-----GMVCKPSKSVHLTVLWRRCRRGGRQCGWIPQIPIISECKSC 232
Db	578	ARSCDSPPRCGGDLCH-GPAIIHA-----NCSRNQ---AWTP--WSSWALCSTSC 632

RESULT 13  
US-09-854-845-4  
: Sequence 4, Application US/09854845

```

1 Patent No. US20020098491A1
2
3 GENERAL INFORMATION:
4
5 APPLICANT: Walke, D. Wade
6 APPLICANT: Wang, Xiaoming
7 APPLICANT: Scoville, John
8 APPLICANT: Turner, C. Alexander Jr.
9
10 TITLE OF INVENTION: No. US20020098491A1el
11
12 FILE REFERENCE: LEX-0177-USA
13
14 CURRENT APPLICATION NUMBER: US/09/854,845
15 CURRENT FILING DATE: 2001-05-14
16
17 PRIOR APPLICATION NUMBER: US 60/205,274
18 PRIOR FILING DATE: 2000-05-18
19
20 PRIOR APPLICATION NUMBER: US 60/208,893
21 PRIOR FILING DATE: 2000-06-02
22
23 NUMBER OF SEQ ID NOS: 50
24
25 SOFTWARE: FastSeq for Windows Version 4.0
26

```

```

; SEQ ID NO 4
; LENGTH: 1093
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-4

```

Query Match	6.8;	Score 86.5;	DB 10;	Length 1093;
Best Local Similarity	30.2%;	Pred. No. 6.7;		
Matches 35;	Conservative 10;	Mismatches 44;	Indels 27;	Gaps 9;

Qy	132	KKQRLSKLRRKLQWLWSQ---TFCPLVYANDLGSRTWPRVY-----KVGSFCBK	181
Db	519	KKQRLSKLRRKLQWLWSQ---TFCPLVYANDLGSRTWPRVY-----KVGSFCBK	181
Qy	182	RSQVPE-----GMVKPKSKSVHLTVLRRCQRGGQRGQWPIQPIIIECKKSC	232
Db	578	ARSCDPRRCGGGLDCL-GRPIHIA-----NCSRNG-----AWTP-WSSWALGSTSC	632

RESULT 14  
US-09-854-845-12  
; Sequence 12, Application US/09854845

```

; Patent NO.: US20020098491A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Walke, D. Wade
; APPLICANT: Wang, Xiaoming
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
;
; TITLE OF INVENTION: No. US20020098491A1 Human Semaphorin Homologs and Polynucleotid
;
; FILE REFERENCE: LEX-0177-USA
;
; CURRENT APPLICATION NUMBER: US/09/854, 845

```

```

; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/205,274
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/208,893
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1136
; TYPE: prt
; ORGANISM: homo sapiens
; US-09-854-845-12

```

Query Match	6.8%	Score 86.5;	DB 10;	Length 1136;
Best Local Similarity	30.2%;	Pred. No. 7;		
Matches 35;	Conservative 10;	Mismatches 44;	Indels 27;	Gaps 9;

Qy	132	KKQRLSKLRRKLQWLMSQ--TFCVLYANNDLGSRFWPRVY-----KVGSFCFK 181
Db	577	KKQRC--TLEDSSNLSLTQNTITACFVRNTRVDGGFGPWSQPCHELDGNSGSLCR 635
Qy	182	-RSCSYPE----GMVCKPSKSVHLTVLRWRCRRGGQRCCGWIPIQPIISECKKSC 232
Db	636	ARSCDSFRRCGGSLDCL-CPATHIA---NCSRNG---AWTF--WSSWALCSTSC 690

RESULT 15  
US-09-854-845-10  
; Sequence 10, Application US/09854845

```

/ Patent No. US20020098491A1
/
/ GENERAL INFORMATION:
/
/ APPLICANT: Wang, D. Wade
/ APPLICANT: Wang, Xiaoming
/ APPLICANT: Scoville, John
/ APPLICANT: Turner, C. Alexander Jr.
/
/ TITLE OF INVENTION: No. US20020098491A1a1el
/
/ FILE REFERENCE: LEX-0177-USA
/
/ CURRENT APPLICATION NUMBER: US/09/854,845
/
/ CURRENT FILING DATE: 2001-05-14
/
/ PRIOR APPLICATION NUMBER: US 60/205,274
/
/ PRIOR FILING DATE: 2000-05-18
/
/ PRIOR APPLICATION NUMBER: US 60/208,893
/
/ PRIOR FILING DATE: 2000-06-02
/
/ NUMBER OF SEQ ID NOS: 50
/
/ SOFTWARE: FastSeq for Windows Version 4.0
/
/ SEQ ID NO 10
/
/ LENGTH: 1151

```

```

; TYPE: PRT
; ORGANISM: homo sapiens
US-09-854-845-10

```

Query Match	6.8%;	Score 86.5;	DB 10;	Length 1151;
Best Local Similarity	30.2%;	Pred. No. 7.1;		
Matches 35;	Conservative 10;	Mismatches 44;	Indels 27;	Gaps 9;

QY	132	KKQRLSKKLRLKQWLMSQ---TFCPVLYANNDLGSRFWPYV-----KVGSFSPK	181
DB	577	KKQRCRCS---TLEDSSNNLSLMTQITACFVRNVTDDGGFGPWSQPQCEHLDGNSGSLCLR	635
QY	182	-RSCSVPE---GMVCKPKSKVHLTVLRWRCRRGQRGQRGWIPIQYPIISECKKSC	232
DB	636	ARSCDSFRPRCGGLDCL-GPAITHIA---NCSRNG---AWTF---WSSWALGSTSC	680

Search completed: January 7, 2003, 09:28:32  
Job time : 12 secs

; APPLICANT: Turner, C. Alexander Jr.  
 ; TITLE OF INVENTION: No. US20020098491A1el Human Semaphorin Homologs and Polynucleotid  
 ; FILE REFERENCE: LEX-0177-USA  
 ; CURRENT APPLICATION NUMBER: US/09/854, 845

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2003, 09:22:07 ; Search time 33 Seconds  
(without alignments)  
1448.574 Million cell updates/sec

Title: US-09-897-322-2  
Perfect score: 1268  
Sequence: 1 MERCPSLGVTLYALVVVLGL.....QRCGWIPQYPIISECKKSC 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	102	8.0	2393	2 Q9Z5F4	Q9Z5F4 myxococcus
2	99.5	7.8	413	4 Q9UBU6	Q9UBU6 homo sapien
3	96	7.6	507	10 Q93VF7	Q93VF7 arabidopsis
4	94.5	7.5	684	2 Q44055	Q44055 aeromonas h
5	93.5	7.4	322	16 Q9RK51	Q9RK51 streptomyce
6	93.5	7.4	372	13 Q12970	Q12970 gallus gall
7	93.5	7.4	372	13 Q12972	Q12972 gallus gall
8	92	7.3	538	15 Q9WJP4	Q9WJP4 moloney mur
9	92	7.3	1737	15 Q92808	Q92808 moloney mur
10	90	7.1	142	10 Q42322	Q42322 arabidopsis
11	88.5	7.0	931	12 Q9QAP6	Q9QAP6 rangiferine
12	88.5	7.0	979	4 Q8TD19	Q8TD19 homo sapien
13	88.5	7.0	979	4 Q8TCY4	Q8TCY4 homo sapien
14	88	6.9	278	16 Q8UGW7	Q8UGW7 agrobacteri
15	88	6.9	814	2 Q59129	Q59129 arthrobacte
16	88	6.9	816	2 Q93NH5	Q93NH5 arthrobacte

17	87.5	6.9	521	2 Q9KI55	Q9KI55 pseudomonas
18	87.5	6.9	728	4 Q9UPH9	Q9UPH9 homo sapien
19	87.5	6.9	1386	5 Q9NKKJ2	Q9NKKJ2 leishmania
20	87	6.9	1975	5 Q9VCD1	Q9VCD1 drosophila
21	86.5	6.8	351	10 Q94GL5	Q94GL5 oryza sativ
22	86.5	6.8	420	10 Q9ZSY5	Q9ZSY5 arabidopsis
23	86.5	6.8	457	16 Q9RRJ4	Q9RRJ4 deinococcus
24	86.5	6.8	529	2 P94909	P94909 microbacter
25	86.5	6.8	622	10 Q9ZSY6	Q9ZSY6 arabidopsis
26	86.5	6.8	622	10 Q9C6B8	Q9C6B8 arabidopsis
27	86.5	6.8	1202	4 Q9P283	Q9P283 homo sapien
28	86	6.8	525	10 P93520	P93520 zea mays (m
29	86	6.8	981	15 Q92809	Q92809 abelson mur
30	85	6.7	599	10 Q41791	Q41791 zea mays (m
31	85	6.7	1621	4 Q13395	Q13395 homo sapien
32	85	6.7	5146	6 Q8SPM4	Q8SPM4 bos taurus
33	84.5	6.7	212	13 Q9W7C5	Q9W7C5 gallus gall
34	84	6.6	363	17 Q9YAX6	Q9YAX6 aeropyrum p
35	84	6.6	442	13 Q9DFG7	Q9DFG7 brachydanio
36	84	6.6	640	10 Q96397	Q96397 chlamydomon
37	83	6.5	266	5 Q18862	Q18862 caenorhabdi
38	83	6.5	476	12 Q80890	Q80890 herpesvirus
39	83	6.5	538	15 Q56222	Q56222 moloney mur
40	83	6.5	1066	10 Q22804	Q22804 arabidopsis
41	83	6.5	2111	3 Q9C102	Q9C102 schizosacch
42	82	6.5	122	16 Q9RSP2	Q9RSP2 deinococcus
43	82	6.5	339	4 Q00320	Q00320 homo sapien
44	82	6.5	344	16 Q9RZM1	Q9RZM1 deinococcus
45	82	6.5	368	13 Q90853	Q90853 gallus gall

ALIGNMENTS

RESULT 1

Q9Z5F4	PRELIMINARY;	PRT; 2393 AA.
ID	Q9Z5F4	
AC	Q9Z5F4;	
DT	01-MAY-1999 (Tremblrel. 10, Created)	
DT	01-MAY-1999 (Tremblrel. 10, Last sequence update)	
DT	01-MAY-2002 (Tremblrel. 20, Last annotation update)	
DE	Tal (Fragment).	
GN	TAL.	
OS	Myxococcus xanthus.	
OC	Bacteria; Proteobacteria; delta subdivision; Myxobacteria;	
OC	Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.	
OX	NCBI_TaxID=34;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ER-15;	
RX	MEDLINE=99141200; PubMed=9973564;	
RA	Paitan Y., Alon G., Orr E., Ron E.Z., Rosenberg E.;	
RT	"The first gene in the biosynthesis of the polyketide antibiotic TA of	
RT	Myxococcus xanthus codes for a unique PKS module coupled to a peptide	
RT	synthetase.";	
RL	J. Mol. Biol. 286:465-474(1999).	
DR	EMBL; AJ006977; CAB38084.1; .	
DR	HSSP; P14687; 1AMU.	
DR	InterPro; IPR000873; AMP-bind.	
DR	InterPro; IPR001242; Condensatn.	
DR	InterPro; IPR000794; Ketoacyl-synt.	
DR	InterPro; IPR003880; Ppanthe attach.	
DR	Pfam; PF00501; AMP-binding; 1.	
DR	Pfam; PF00668; Condensation; 1.	
DR	Pfam; PF00109; ketoacyl-synt; 2.	
DR	Pfam; PF02801; ketoacyl-synt_C; 1.	
DR	Pfam; PF00550; pp-binding; 2.	
DR	PRINTS; PR00154; AMPBINDING.	
DR	PROSITE; PS00075; ACP DOMAIN; 2.	
DR	PROSITE; PS00455; AMP_BINDING; 1.	
FT	NON_TER 1	
FT	NON_TER 2393	
SQ	SEQUENCE 2393 AA; 259537 MW; 6CF6E0153BC3E126 CRC64;	

Query Match 8.0%; Score 102; DB 2; Length 2393;  
 Best Local Similarity 23.7%; Pred. No. 1.6;  
 Matches 56; Conservative 27; Mismatches 67; Indels 86; Gaps 12;

7 LGVTLVAVVGLRATPAGQH-----YLHTRPAPSDNLPVLDLIEHPDIFDEREK 59  
 DB LATAICGFNVMAVARAGL-GQHSFGLRLH-----DS--VIDGLEHMYFFPRVVK 294

QY 60 DINEITLRLGLGHYPGF-----MATSPEDRPGGGGAGAEADLADQLR 109  
 DB 295 D-----LRLSNGPEEAPGFQMTFTFOSLQLTSAAPRPEPSGG-----LPLEPLDC 341

QY 110 QRPSSGMPSEIKGLEFSEGLAQKK-----ORLSKLRKLM----- 147  
 DB 342 VHOEGAYPLEVEVEGKGLTLHKYDARLYEADTVERMKRLRAADQYADVESPUSA 401

QY 148 --WL-----WSQFCPVLYAMNDLSRFWPRVYKVGSCFSKSGSVPEGM 190  
 DB 402 LSWLDEERRTLLRDMNATATPL---EDLG-----VHELFGQARETPDAM 445

RESULT 2  
 QY 09UB6 PRELIMINARY; PRT; 413 AA.  
 AC QY 09UB6:  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE Autosomal highly conserved protein.  
 NS AHCP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Bourgeon T., Jamin S., Leroy P., McElreavey K., Fellous M.;  
 RT "Identification of AHCP, a novel highly conserved gene on human  
 RT chromosome 6 candidate for genetic predisposition to schizophrenia.";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF097027; AAF07850.1; -;  
 DR EMBL; AF097026; AAF07849.1; -;  
 SQ SEQUENCE 413 AA; 44123 MW; 35B92497907CCD2E CRC64;

Query Match 7.8%; Score 99.5; DB 4; Length 413;  
 Best Local Similarity 27.6%; Pred. No. 0.32;  
 Matches 50; Conservative 17; Mismatches 69; Indels 45; Gaps 11;

QY 26 GGOHYIHRPAPSDNLPVLDLIEHP--DRIFDPRKDLNLTLSLLGHHYDPGFMAISP 83  
 DB 19 GGDH-----BPVPLRGPPTTAVPCPRDDPAEPOA-----PG----- 51

QY 84 PDRPGGGAAGAEADLADQLROR-----PSGAMPEIKGLEFSEGLA-QGKQRL 137  
 DB 52 ---RPAPELAAAAADKLEPRELKRGAAGSGAELEQACCEAPLAAPRERARL 108

QY 138 S-KLRLKRLQMWLSQTFCEPVLYAMNDLSRFWPRVYK---VSGSCFSKSGSVPEGVCK 193  
 DB 109 SAREYSRQVHEWLM-QSYCGYL-TWHS-GLAAPPAYCSPSPSPSGAANAQAAPP 165

QY 194 P 194  
 DB 166 P 166

RESULT 3  
 QY 093VF7 PRELIMINARY; PRT; 507 AA.  
 AC QY 093VF7:  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Putative WD-40 repeat protein, MS14 (At2g19520/F3P11.12).  
 GN AT2G19520 OR F3P11.12.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OC NCBI\_Taxid=3702;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,  
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R.W.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RL "Arabidopsis cDNA clones."  
 RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,  
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R.W.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AY059799; AAL24281.1; -;  
 DR EMBL; AY057655; AAL15286.1; -;  
 DR EMBL; AY081447; AAM10009.1; -;  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 6.  
 DR Prodom; PD000018; WD40; 1.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; 3.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
 DR Repeat; WD repeat.  
 SQ SEQUENCE 507 AA; 55759 MW; C37F8000F8B33397 CRC64;

Query Match 7.6%; Score 96; DB 10; Length 507;  
 Best Local Similarity 30.5%; Pred. No. 0.89;  
 Matches 29; Conservative 10; Mismatches 40; Indels 16; Gaps 4;

QY 80 ATSPEDRPGGGAAGAEADLADQLRORPSGAMPEIKGLEFSEGLAQKKQ 135  
 DB 8 AVSPQATTSFGGCGAGGPKKRGKPKTKEDS---QTPSSQGSQSVKME-----SGKKT 58

QY 136 RLKSLKRLKRLQMWLSQTFCEPVLYAMNDLSRFWPRVYK 170  
 DB 59 QQSPSVDEKYSQW---KGLVPLIYDILANHLNLPW 90

RESULT 4  
 QY 044055 PRELIMINARY; PRT; 684 AA.  
 AC Q44055:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Phospholipase C.  
 OS Aeromonas hydrophila.  
 OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
 OC Aeromonas.

```
OX NCBI_TaxID=644;
RN [1]
RP SEQUENCE OF 1-619 FROM N.A.
RC STRAIN=JMP636;
RX MEDLINE=95284718; PubMed=7767226;
RA Ingham A.B.; Pemberton J.M.;
RT "A lipase of Aeromonas hydrophila showing nonhemolytic phospholipase C
activity.";
RL Curr. Microbiol. 31:28-33(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JMP636;
RA Ingham A.B.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U14011; AAA75598.1; -.
DR InterPro; IPR000734; Lipase.
DR PROSITE; PS00120; LIPASE_SER; 1.
SQ SEQUENCE 684 AA; 72396 MW; 702745900A9A5714 CRC64;

Query Match 7.5%; Score 94.5; DB 2; Length 684;
Best Local Similarity 25.2%; Pred. No. 1.8;
Matches 58; Conservative 22; Mismatches 89; Indels 61; Gaps 10;

QY 19 GLRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDPKEKDLNETLLSLGG---HYD 75
Db 289 GSRVAPASGNKDA-IATAVAPSRPLAELRPRFIPGRSSCLTSSPRLAPLVPGIKPRPS 347
QY 76 PGFMATSP-----PEDRPGG---GGG-----AAGGAEDLAELQ 106
Db 348 PGTVPSQPVHRQCAGVTVSDRRPGRRCGSLANGTDCRSDSPRAAGGGEQAHRGDP 407
QY 107 LLRQRSPGAMPSEIKLEFSEGLAQKKORLSKKLRKQLMWLWSOTFCPVLYAMNDL-- 164
Db 408 HLGRQAGSR-AEYRPLQPAAGRAIGRAIDAHLCRRRAV-----CHHRCDLHLSAHHLGE 462
QY 165 GSRFWPRYKVGSCFSKRSKSPVEGVMCKPSKVHLTVLRWCQRGGQR 214
Db 463 GERLCP-----GAGPDRCRCSKHGQEC-----RRGGDR 491

RESULT 5
Q9RK51 PRELIMINARY; PRT; 322 AA.
ID Q9RK51
AC Q9RK51;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DE Hypothetical protein SC00326.
GN SC00326 OR SCF12.05.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D.; Chater K.F.; Cerdano-Tarraga A.-M.; Challis G.L.;
RA Thomson N.R.; James K.D.; Harris D.E.; Quail M.A.; Kieser H.;
RA Harper D.; Bateman A.; Brown S.; Chandra G.; Chen C.W.; Collins M.;
RA Cronin A.; Fraser A.; Goble A.; Hidalgo J.; Hornsby T.; Howarth S.;
RA Huang C.-H.; Kieser T.; Larke L.; Murphy L.; Oliver K.; O'Neill S.;
RA Rabinowitz E.; Rajadream M.A.; Rutherford K.; Rutter S.;
RA Seeger K.; Saunders D.; Sharp S.; Squares S.; Taylor K.;
RA Warren T.; Wietzorrek A.; Woodward J.; Barrell B.G.; Parkhill J.;
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL117669; CAB56131.1; -.
KW Hypothetical protein.
SQ SEQUENCE 322 AA; 35339 MW; DD55BB0480090638 CRC64;

Query Match 7.4%; Score 93.5; DB 16; Length 322;
```

```
Best Local Similarity 26.3%; Pred. No. 0.88;
Matches 45; Conservative 21; Mismatches 50; Indels 55; Gaps 10;

QY 9 VTLVYALVVLGLRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPDP--KEKDLNETLL 66
Db 131 LAVHVLVTAAGVRVTP-----HFVGAEPARTGN-----GPRGLRPLGPEDLARELI 178
QY 67 RSLGGHYDGFEMATSPPEDRPGGGGAGGAEDLAELDQLLRQRP---SGAMPSEIKGL 123
Db 179 TSLGGTARASALFAATPPD-----DILTRADPFADPGLLP---EGL 216
QY 124 EFSSEGLAQKKORLSKKLRKQLMWLWSOTFCPVLYA---WND-----LGS 166
Db 217 RY-DRLAPQQQLLERLVGRYLDLDR-----APEAYARECNSDAVRRGLGS 259

RESULT 6
O12970 PRELIMINARY; PRT; 372 AA.
ID O12970
AC O12970;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE Radical fringe.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIME;
RX MEDLINE=97238679; PubMed=9121551;
RA Rodriguez-Esteban C.; Schwabe J.W.; De La Pena J.; Foys B.;
RA Eshelman B.; Belmonte J.C.;
RT "Radical fringe positions the apical ectodermal ridge at the
dorsoventral boundary of the vertebrate limb.";
RL Nature 386:360-366(1997).
DR EMBL; U82088; AAC60107.1; -.
DR InterPro; IPR003378; Fringe.
DR Pfam; PF02434; Fringe; 1.
SQ SEQUENCE 372 AA; 40962 MW; 84CB7B74A4B81C6D CRC64;

Query Match 7.4%; Score 93.5; DB 13; Length 372;
Best Local Similarity 25.4%; Pred. No. 1.1;
Matches 52; Conservative 17; Mismatches 81; Indels 55; Gaps 11;

QY 4 CPSLGVTLVYLVGLRATPAGGQHYLHIRPAPSDNLPVLDLIEHPDIPD--PKEKDL 61
Db 12 CFFLSVTAALVLLLLPRGQPPAAPRR---RPPAPG-----SRPSPKREARPAAGSDV 61
QY 62 NETLRLSLGGHYDGFEMATSPPEDRPGGGGAGGAEDLAEL---DQLLRQRPSCAMPS 118
Db 62 -----PGDR-GGGGAGGGRGVAGSPWPSRRVRMGPPGGSAK 98
QY 119 EIKGLEFSEGLAQKKORLSKKLRKQLMWLW-----SOTFCPVLYAMNDLGSRFWRPV 173
Db 99 E--SLEKIDIFATKTRKYKHTRLELLPQTWISRAGQTF--IFTDWEDELRL-----L 149
QY 174 KVGSCFSKRSKSCS---VPEGMVCKPS 195
Db 150 KAGDHMINTCNSAVHTRQALCKMS 174

RESULT 7
O12972 PRELIMINARY; PRT; 372 AA.
ID O12972
AC O12972;
DT 01-JUL-1997 (TREMELrel. 04, Created)
DT 01-JUL-1997 (TREMELrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE Radical fringe.
OS Gallus gallus (Chicken).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97238680; PubMed=9121552;  
 RA Lauter E., Dahn R., Orozco O.E., Yeo C.Y., Piseni J., Henrique D.,  
 RA Abbot U.K., Fallon J.F., Tabin C.;  
 RT "Expression of Radical Fringe in limb-bud ectoderm regulates apical  
 RT ectodermal ridge formation.";  
 RL Nature 386:366-373(1997).  
 DR EMBL: U91850; AAC60100.1; -;  
 DR InterPro: IPR003378; Fringe.  
 DR Pfam: PF02434; Fringe; 1.  
 SQ SEQUENCE 372 AA; 40904 MW; 05BA18254EA079D6 CRC64;  
 Query Match 7.4%; Score 93.5; DB 13; Length 372;  
 Best Local Similarity 25.4%; Pred. No. 1.1;  
 Matches 52; Conservative 17; Mismatches 81; Indels 55; Gaps 11;  
 QY 4 CSLSGVTYALVAVVGLRATPAGQHYLIRPAPSDNLPLVDLIEHDPDIFD--PKRKDL 61  
 DB 12 CFLSTVTAVALLLLLPRGQPPAPRR--RPPAPG-----SRSPKREAPAGSDV 61  
 QY 62 NETLRLSLGGHYDPGFMTSPEDRPGGGGAAGADLAEL---DQLRKQPSGAMS 118  
 DB 62 -----PQDR-GGGSGAAGGGRGVAAGSPMRPRRYMGPSPGSAK 98  
 QY 119 EIKGLFSEGLAQCKKQRLSKLRKLQMWLW-----SQFPCPVLYAMNDIGSRF 173  
 DB 99 E-SLEKLDFFIVKTRKXHKTRLELLFQTWISRAQGTF--IFTMEDRELK-----L 149  
 QY 174 KVSGSCFSKRSKRS---VPEGWCKPS 195  
 DB 150 KAGDHMINTCASAVHTRQALCCWKS 174  
 RESULT 8  
 Q9WJ2P4 PRELIMINARY; PRT; 538 AA.  
 AC Q9WJ2P4;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Pr65.  
 GN GAG.  
 OS Moloney murine leukemia virus.  
 OC Viruses; Retrovirdae; Retroviridae; Gammaretrovirus.  
 OK NCBI\_TaxID=11801;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Petropoulos C.J.;  
 RT "Appendix 2 - Retroviral taxonomy, protein structure, sequences, and  
 RT genetic maps.";  
 RL (In) Coffin J.M. (eds.);  
 RL Retroviruses, pp.757-757, Cold Spring Harbor Laboratory Press,  
 RL New York (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Chappey C.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF033811; AAC82566.1; -;  
 DR InterPro: IPR000840; Gag\_MA.  
 DR InterPro: IPR002079; Gag\_p12.  
 DR InterPro: IPR003036; Gag\_p30.  
 DR InterPro: IPR002955; P\_titch\_extensn.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF01140; Gag\_MA; 2.  
 DR Pfam: PF01141; Gag\_p12; 2.  
 DR Pfam: PF02093; Gag\_p30; 2.  
 DR Pfam: PF00098; zf-CCHC; 2.  
 DR PRINTS: PRO1217; PRICHEXTENSN.

DR SMART: SM00343; Znf\_C2HC; 2.  
 KM Core protein; Polypeptin.  
 FT CHAIN 2 131 P15 MA.  
 FT CHAIN 132 215 P12.  
 FT CHAIN 216 478 P10 CA.  
 FT CHAIN 479 534 P10 NC.  
 FT CHAIN 535 538 P14 PR.  
 SQ SEQUENCE 538 AA; 60858 MW; 8A7652439B464495 CRC64;  
 Query Match 7.3%; Score 92; DB 15; Length 538;  
 Best Local Similarity 26.5%; Pred. No. 2.3;  
 Matches 45; Conservative 12; Mismatches 45; Indels 68; Gaps 9;  
 QY 21 RAPPAGQHY-----LHTRAP---SDN-LPLVDLI-EHPPIPIPKKDKDNTLRLSL 69  
 DB 122 RSTPPRSSLYPALTPSLGAPKPOVLSDSGPLDILDLPDPPYRDP----- 169  
 QY 70 LGGHYDPGFMTSPEDRPGGGGA--AGADLAELDQLRQ-----PSGAMS 118  
 DB 170 -----PPSDRONGGEATPAGAPDPSPMASRLAGRRPPVADSTTQAPFL 217  
 QY 119 EIKGLFSEGLAQCKKQRLSKLRKLQMWLSQFPCPVLYAMNDIGSRF 168  
 DB 218 RAGG-----NGQLQYWPFS--DLVWKNKNNPSP 245  
 RESULT 9  
 Q92808 PRELIMINARY; PRT; 1737 AA.  
 AC Q92808;  
 DT 01-NOV-1998 (TREMblrel. 08, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Pr180.  
 GN GAG-POL.  
 OS Moloney murine leukemia virus.  
 OC Viruses; Retrovirdae; Retroviridae; Gammaretrovirus.  
 OK NCBI\_TaxID=11801;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Petropoulos C.J.;  
 RT "Appendix 2 - Retroviral taxonomy, protein structure, sequences, and  
 RT genetic maps.";  
 RL (In) Coffin J.M. (eds.);  
 RL Retroviruses, pp.757-757, Cold Spring Harbor Laboratory Press,  
 RL New York (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Chappey C.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 CC KNOWN AS THE RETROPEPSIN FAMILY.  
 DR EMBL: AF033811; AAC82566.1; -;  
 DR HSP: P03355; IMML.  
 DR InterPro: IPR001995; Aspprotease\_rtrv.  
 DR InterPro: IPR001969; Aspprotease\_bite.  
 DR InterPro: IPR000840; Gag\_MA.  
 DR InterPro: IPR002079; Gag\_p12.  
 DR InterPro: IPR003036; Gag\_p30.  
 DR InterPro: IPR002156; RNaseH.  
 DR InterPro: IPR001584; Rve.  
 DR InterPro: IPR00477; RVTbe.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR Pfam: PF01140; Gag\_MA; 2.  
 DR Pfam: PF01141; Gag\_p12; 2.  
 DR Pfam: PF02093; Gag\_p30; 2.  
 DR Pfam: PR00075; rnaseh; 2.  
 DR Pfam: PF00665; rve; 2.  
 DR Pfam: PF00077; rvp; 2.  
 DR Pfam: PF00078; rvc; 1.  
 DR Pfam: PF00098; zf-CCHC; 2.  
 DR SMART: SM00343; Znf\_C2HC; 2.  
 DR PROSITE: PS00141; ASP\_PROTEASE; 2.



```

Db 625 KCGQLGVNKKRLGILNGLGPIGKQVIRVSCGDEFTIATDDNHLFAMNGNGRLAM 684
Qy 34 ----RPAQSDNLEPLVDLIEHPDPIFDP-----KENDLNETHLRLSLGG 72
Db 685 PTERPHGSD-----ICTSWPRPIFGSLHHVPLDSCRGMHTLILVEKVLMSKTRNSSG 739
Qy 73 -HYDPGFMAITSPEDRPGGGGAAGAEADLAELDQLRQPSGAMPEIKGLEPSEGL-- 129
Db 740 LSIQVIFGSSS-----PGGGGGGGGGGEEEDSQDESETPD--PSGGFRGTMEADRMGLIS 793
Qy 130 ---AAGKQRLSKRLRKLQMWLMSQTCFVLYAMNDLGRF 168
Db 794 PTEAMGNSNGASSSCPGWLKLENAEFIMPDPSPSLSAF 835

RESULT 13
O8TCY4 PRELIMINARY; PRT; 979 AA.
AC O8TCY4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE NIMA-family kinase NERCC.
GN NERCC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rold J., Mikhailov A., Belham C., Avruch J.;
RT "Nercc, a mammalian NIMA-family kinase, binds the Ran GTPase and
RT regulates mitotic progression."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY080896; AAL87410.1; -.
KW Kinase.
SQ SEQUENCE 979 AA; 107034 MW; 002483C1711BADA CRC64;

Query Match 7.0%; Score 88.5; DB 4; Length 979;
Best Local Similarity 22.5%; Pred. No. 10;
Matches 50; Conservative 21; Mismatches 84; Indels 67; Gaps 7;

Qy 3 RCPSSGVTLYALVVLGLRATPAGGCHYH----- 33
Db 625 KCGQLGVNKKRLGILNGLGPIGKQVIRVSCGDEFTIATDDNHLFAMNGNGRLAM 684
Qy 34 ----RPAQSDNLEPLVDLIEHPDPIFDP-----KENDLNETHLRLSLGG 72
Db 685 PTERPHGSD-----ICTSWPRPIFGSLHHVPLDSCRGMHTLILVEKVLMSKTRNSSG 739
Qy 73 -HYDPGFMAITSPEDRPGGGGAAGAEADLAELDQLRQPSGAMPEIKGLEPSEGL-- 129
Db 740 LSIQVIFGSSS-----PGGGGGGGGGGEEEDSQDESETPD--PSGGFRGTMEADRMGLIS 793
Qy 130 ---AAGKQRLSKRLRKLQMWLMSQTCFVLYAMNDLGRF 168
Db 794 PTEAMGNSNGASSSCPGWLKLENAEFIMPDPSPSLSAF 835

RESULT 14
O8UCW7 PRELIMINARY; PRT; 278 AA.
AC O8UCW7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein Atu0918.
GN ATU0918 OR AGR_C1672.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Seubai J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Bisen J.A., Karp P.D., Boyce D. St.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland B., Palmieri A.,
RA Raymond C., Rouse G., Ssephimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krepan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58 ";
RN Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Guorillo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,
RA Woliam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58 ";
RN Science 294:2323-2328 (2001).
RN [3]
DR EMBL; AE009057; AAL41932.1; -.
DR EMBL; AE008022; AAK86722.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 278 AA; 30704 MW; 73100D0C5F633D4 CRC64;

Query Match 5.9%; Score 88; DB 16; Length 278;
Best Local Similarity 27.9%; Pred. No. 2.5;
Matches 43; Conservative 19; Mismatches 74; Indels 18; Gaps 6;

Qy 16 VWGLRATPAGGCHYHIRPAPSDNLEPLVDLIEHPDPIFDPK--DLNETHLRLSLGGHY 74
Db 104 VVEIDETKQPARNINTAPLTS---LSRLKERSDASAFPPEDALAAGRLAADPHRGL 159
Qy 75 DPGFMATSPEDRPGGGGAAGAEADLAELDQLR--QRPSGAMPEIKGLE-----FS 126
Db 160 NRPVATWMEPR-IASRKGEGAGALDTEAAMARARFRAADAMGPELGVADICPE 218
Qy 127 EGIAGKKQRL---SKRLRKLQMWLMSQTCF 156
Db 219 KGLETERERLWPARSAKLLRLAALSLARHYAP 252

```

```

RESULT 15
O59129 PRELIMINARY; PRT; 814 AA.
AC O59129;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Nicotinic dehydrogenase, large chain (EC 1.5.99.4) (NDHC).
GN NDHC.
OS Arabidopsis thaliana.
OC Plantae; Equisetopsida; Gymnosperms; Angiosperms; Malvales;
OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=29320;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95115562; PubMed=7815950;
RA Grether-Beck S., Igloi G., Pust S., Schilz E., Decker K., Brandach R.;
RT "Structural analysis and molybdenum-dependent expression of the pAO1-
RT encoded nicotinic dehydrogenase genes of Arthrobacter nicotovorans";
RL Mol. Microbiol. 13:929-936 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Brandach R.;

```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2003, 09:19:32 ; Search time 11 Seconds  
(without alignments)  
874.773 Million cell updates/sec

Title: US-09-897-322-2  
Perfect score: 1268  
Sequence: 1 MERCPSLGVTLYALVVVLGL.....QRCGWIPQYPIISECKSC 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1268	100.0	232	1	NOGG_HUMAN	Q13253 homo sapien
2	1258	99.2	232	1	NOGG_MOUSE	P97466 mus musculus
3	1010	79.7	223	1	NOGG_CHICK	O93525 gallus gall
4	949.5	74.9	222	1	NOGG_XENLA	P49011 xenopus lae
5	782	61.7	144	1	NOGG_RAT	Q62809 rattus norv
6	759.5	59.9	223	1	NOG3_BRARE	O9ybv3 brachydanio
7	700.5	55.2	216	1	NOG1_BRARE	Q9w741 brachydanio
8	655	51.7	223	1	NOGG_FUGRU	O9yht8 fugu rubrip
9	546	43.1	212	1	NOG2_BRARE	Q9w740 brachydanio
10	96	7.6	507	1	MSI4_ARATH	O22607 arabidopsis
11	92	7.3	538	1	GAG_MLVMO	P03332 moloney mur
12	89	7.0	1093	1	SM5B_MOUSE	O60519 mus musculus
13	88	6.9	917	1	VGLB_HSVB2	P12841 bovine herp
14	87.5	6.9	226	1	TRUN_DROME	Q24155 drosophila
15	87.5	6.9	722	1	P85B_RAT	O63788 rattus norv
16	87.5	6.9	728	1	P85B_HUMAN	O00459 homo sapien
17	86	6.8	724	1	P85B_BOVIN	P23726 bos taurus
18	85.5	6.7	315	1	PIX1_MOUSE	P70314 mus musculus
19	85.5	6.7	331	1	SHX2_MOUSE	P70390 mus musculus
20	85	6.7	567	1	SSPO_BOVIN	P98167 bos taurus
21	83	6.5	538	1	GAG_MSMVO	P03334 moloney mur
22	82.5	6.5	235	1	GAG_MSVAB	P03333 abelson mur
23	82	6.5	468	1	GAG_MSVMT	P32594 moloney mur
24	80	6.3	331	1	SHX2_HUMAN	O60902 homo sapien
25	79.5	6.3	397	1	HKL2_MALDO	O04135 malus domes
26	79.5	6.3	398	1	HKL1_MALDO	O41334 malus domes
27	79.5	6.3	1310	1	ICP4_VZVD	P09310 varicella-z
28	79	6.2	3511	1	MY15_MOUSE	O9gz24 mus musculus
29	78.5	6.2	722	1	P85B_MOUSE	O08908 mus musculus
30	78	6.2	1269	1	FLIH_HUMAN	Q13045 homo sapien
31	78	6.2	1572	1	BA12_HUMAN	O60241 homo sapien
32	77.5	6.1	388	1	SOX7_HUMAN	O9b581 homo sapien
33	77.5	6.1	490	1	CE05_ECOLI	Q47500 escherichia

RESULT 1									
NOGG_HUMAN	112892	seqs	41476328	residues					
AC	Q13253				STANDARD;	PRT;	232	AA.	
DT	16-OCT-2001	(Rel. 40, Created)							
DT	16-OCT-2001	(Rel. 40, Last sequence update)							
DT	16-OCT-2001	(Rel. 40, Last annotation update)							
DE	Noggin precursor.								
GN	NOG.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Placenta, and Temporal cortex;								
RX	MEDLINE=95359592; PubMed=7666191;								
RA	Valenzuela D.M., Economides A.N., Rojas E., Lamb T.M., Nunez L.,								
RA	Jones P., Ip N.Y., Espinosa R. III, Brannan C.I., Gilbert D.J.,								
RA	Copeland N.G., Jenkins N.A., Le Beau M.M., Harland R.M.,								
RA	Yancopoulos G.D.;								
RT	"Identification of mammalian noggin and its expression in the adult								
RT	nervous system."								
RL	J. Neurosci. 15:6077-6084(1995).								
RN	[2]								
RP	VARIANTS SYM1/SYNS1 R-35; C-189; G-217; N-220; C-222; G-222 AND L-223.								
RX	MEDLINE=99178269; PubMed=10080184;								
RA	Gong Y., Krakow D., Marcelino J., Wilkin D., Chitayat D.,								
RA	Babul-Hirji R., Hudgins L., Cremers C.W., Cremers F.P.M.,								
RA	Brunner H.G., Reinker K., Rimo D.L., Cohn D.H., Goodman F.R.,								
RA	Reardon W., Patton M., Francmano C.A., Warman M.L.;								
RT	"Heterozygous mutations in the gene encoding noggin affect human joint								
RT	morphogenesis."								
RL	Nat. Genet. 21:302-304(1999).								
CC	FUNCTION: ESSENTIAL FOR CARTILAGE MORPHOGENESIS AND JOINT								
CC	FORMATION. INHIBITOR OF BONE MORPHOGENETIC PROTEINS (BMP)								
CC	SIGNALING WHICH IS REQUIRED FOR GROWTH AND PATTERNING OF THE								
CC	NEURAL TUBE AND SOMITE (BY SIMILARITY).								
CC	SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).								
CC	SUBCELLULAR LOCATION: Secreted.								
CC	DISEASE: DEFECTS IN NOG ARE THE CAUSE OF SYMPHALANGISM PROXIMAL								
CC	SYNDROME (SYM1). SYM1 IS CHARACTERIZED BY THE HEREDITARY ABSENCE								
CC	OF THE PROXIMAL INTERPHALANGEAL (PIP) JOINTS (CUSHING								
CC	SYMPHALANGISM). SEVERITY OF PIP JOINT INVOLVEMENT DIMINISHES								
CC	TOWARDS THE RADIAL SIDE. DISTAL INTRAPHALANGEAL JOINTS ARE LESS								
CC	FREQUENTLY INVOLVED AND METACARPOPHALANGEAL JOINTS ARE RARELY								
CC	AFFECTED WHEREAS CARPAL BONE MALFORMATION AND FUSION ARE COMMON.								
CC	IN THE LOWER EXTREMITIES, TARSAL BONE COALITION IS COMMON.								
CC	CONDUCTIVE HEARING LOSS IS SEEN AND IS DUE TO FUSION OF THE STAPES								
CC	TO THE PETROUS PART OF THE TEMPORAL BONE.								
CC	DISEASE: DEFECTS IN NOG ARE THE CAUSE OF MULTIPLE SYNOSTOSES								
CC	SYNDROME 1 (SYNS1). ALSO KNOWN AS SYNOSTOSES, MULTIPLE, WITH								
CC	BRACHYDACTYLIS/SYMPHALANGISM-BRACHYDACTYLIC SYNDROME. SYNS1 IS								
CC	CHARACTERIZED BY TUBULAR-SHAPED (HEMICYLINDRICAL) NOSE WITH LACK								
CC	OF ALAR FLARE, OTOSCLEROTIC DEAFNESS, AND MULTIPLE PROGRESSIVE								
CC	JOINT FUSIONS COMMENCING IN THE HAND. THE JOINT FUSIONS ARE								

Q47125 escherichia  
O64239 mycobacteri  
P28707 saccharomyc  
P78337 homo sapien  
O01970 homo sapien  
P07700 meleagris g  
O14529 homo sapien  
Q96rt7 homo sapien  
P53783 mus musculu  
P26032 measles vir  
O921b8 mus musculu  
P72966 synechocyst

ALIGNMENTS

[illegible]

Ox	Mammalia; Euteria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
Nc31_taxid=10090;	(1)
Rp	SEQUENCE FROM N.A.
Rc	STRAIN=129/Sv;
Ra	MEDLINE=9852829; PubMed=9585504;
Rm	McMahon J.A., Takada S., Zimmerman L.B., Fan C.-M., Harland R.M., McMahon A.P.;
Rt	"Noggin-mediated antagonism of BMP signaling is required for growth RT and patterning of the neural tube and somite.";
Rl	Genes Dev. 12:1438-1452(1998).
[2]	
Rp	FUNCTION.
Ra	MEDLINE=96267313; PubMed=9603738;
Rm	Bruner L.J., McMahon J.A., McMahon A.P., Harland R.M.;
Rt	"Noggin, cartilage morphogenesis, and joint formation in the mammalian skeleton.";
Rl	Science 280:1455-1457(1998).
Cc	-I- FUNCTION: ESSENTIAL FOR CARTILAGE MORPHOGENESIS AND JOINT FORMATION. INHIBITOR OF BONE MORPHOGENETIC PROTEINS (BMP) SIGNALING WHICH IS REQUIRED FOR GROWTH AND PATTERNING OF THE NEURAL TUBE AND SOMITE.
Cc	-I- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
Cc	-I- TISSUE SPECIFICITY: EXPRESSED IN CONDENSING CARTILAGE AND IMMATURE CHONDROCYTES.
Cc	-I- DEVELOPMENTAL STAGE: EMBRYONIC EXPRESSION WAS FIRST DETECTED IN THE NODE AT 7.5 DAYS POSTCOITUM. BY EARLY SOMITE STAGES, EXPRESSION EXTENDS ANTERIORLY ALONG THE ENTIRE LENGTH OF THE NOTOCHORD AND IS EXPRESSED IN THE DORSAL NEURAL TUBE FROM THE CAUDAL HINDERBRAIN TO THE POSTERIOR-MOST REGION OF THE EMBryo. BY THE TIME CRANIAL TUBE CLOSURE IS COMPLETED EXPRESSION IS CONTINUOUS ALONG MOST OF THE DORSAL MIDLINE OF THE NEURAL TUBE, TO ITS POSTRAL TERMINATION AT THE BASE OF THE FOREBRAIN. EXPRESSION IN THE NEURAL TUBE AND CAUDAL NOTOCHORD REMAINS UNCHANGED DURING EARLY ORGANOGENESIS FROM 9.5 DPC TO 10.5 DPC.
Cc	-I- DISEASE: DEFECTS IN NOG ARE THE CAUSE OF A RECESSIVE LETHAL PHENOTYPE AT BIRTH. MULTIPLE DEFECTS INCLUDE A FAILURE OF NEURAL TUBE CLOSURE, BROAD CLUB-SHAPE LIMBS, LOSS OF CAUDAL VERTEBRAE, A SHORTENED BODY AXIS, AND RETENTION OF A SMALL VESTIGIAL TAIL.
Cc	-I- SIMILARITY: BELONGS TO THE NOGGIN FAMILY.
Cc	-----
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch).
Cc	EMBL; U79163; AAB38281.1; -. DR MGD; MGI:104327; Nog. KW Glycoprotein; Signal. FT SIGNAL 1 27 POTENTIAL. FT CHAIN 28 232 NOIGIN. FT CARBOHD 62 62 N-GLUCN (GLCNAC. . .) (POTENTIAL). SQ SEQUENCE 232 AA; 25770 MW; 0192AF6373B78B74 CRC64;
Qy	Query Match 99.2%; Score 1258; DB 1; Length 232; Best Local Similarity 99.1%; Pred. No. 1.4e-105; Matches 230; Conservative 0; Mismatched 2; Indels 0; Gaps 0;
Dd	1 MERCSLGVTYALVAVVGLRATPAAGGGHYHIRAPSDNLPLVDLIHPDIPPKCKD 60
Dd	1 MERCSTLGVTLVALVVVGLRAAPGGGHYLHIRAPSDNLPLVDLIHPDIPPKCKD 60
Qy	61 LNFTLLRSILGGHYDPGMATSPEDRPFGGGGAAGAEDLAELDQLLRORPSGAMPSSEI 120
Dd	61 LNFTLLRSILGGHYDPGMATSPEDRRGGGGGARGAEDLAELDQLLRORPSGAMPSSEI 120
Qy	121 KGLEFSBGLAGOKKRLSKLRRKIOMLWSQTFCPVLYANNLDGSFRWPVYXGSCFS 180
Dd	121 KGLEFSBGLAGOKKRLSKLRRKIOMLWSQTFCPVLYANNLDGSFRWPVYXGSCFS 180





CC GILL ARCHES.  
CC -!- SIMILARITY: BELONGS TO THE NOGGIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF084949; RAD09176.1; -.  
CC ZFIN; ZDB-GENE-990714-8; nog3.  
KW Glycoprotein; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 223 NOGGIN 3.  
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 223 AA; 26029 MW; A21AE5DA36875A37 CRC64;  
  
Query Match 59.4%; Score 759.5; DB 1; Length 223;  
Best Local Similarity 58.4%; Pred. No. 5.7e-61;  
Matches 136; Conservative 36; Mismatches 48; Indels 13; Gaps 4;  
  
QY 1 MERCPSELGVTLYALVVLGLRATPAGQHYLHTRPAPSNLPLVDLIEHPDPIDPKED 60  
DB 1 MDNIPYFLATV--LIFSLGPRIBEGMCQHYLLRPIPSDSLPIVELKEDDPVLDPKED 58  
QY 61 LNETLRSLLGGHYDFGFWATSPEDRPGCGGGAAGDAELDLQRLQRPSPGMPSEI 120  
DB 59 LNETLRAILGSHFEQNFMSINPEDKHAG-----QDELNESELMKQRPNGIMPKEI 110  
QY 121 KGLEFSEGLAOGKQRLSKKRLKQMLWSOTFCPLVYAWNDLGRFWRPVYVKGVCSCFS 180  
DB 111 KAMEFD--IQHGKKHFKPSKLRRLQLWLSYTCFVHVHTQDLGNRFVRYLKVGVSCYN 168  
QY 181 KRSCSVPEGMVCKPSKSVHLTVLRWC-QRRGQRCGWIPYPIIIECKKSC 232  
DB 169 KRSCSVPEGMVCKPPKSSHLTVLRWCVRQKGLKCAWIPVQYPIIIECKKSC 221  
  
RESULT 7  
ID NOGI BRARE STANDARD; PRT; 216 AA.  
AC Q9W741;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Noggin i precursor.  
GN NOGI.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99423658; PubMed=10491267;  
RA Fuerthauer M., Thisse B., Thisse C.;  
RT "Three different noggin genes antagonize the activity of bone  
morphogenetic proteins in the zebrafish embryo.";  
RL Dev. Biol. 214:181-196(1999).  
CC -!- FUNCTION: INHIBITOR OF BONE MORPHOGENETIC PROTEINS (BMP)  
SIGNALING. MAY PLAY AN IMPORTANT ROLE IN THE DORSOVENTRAL  
PATTERNING OF THE EMBRYO.  
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DEVELOPMENTAL STAGE: DETECTED FOLLOWING THE ACTIVATION OF THE  
ZYGOTIC GENOME IN A FEW DEEP CELLS OF THE MARGINAL REGION OF THE  
BLASTODERM. FROM THE 5-12 SOMITE STAGE, EXPRESSION IS OBSERVED IN  
THE DORSAL TELENCEPHALON AND IN POSTERIOR AND VENTRAL PARTS OF THE  
EYE FIELD. BY THE 12-SOMITE STAGE DETECTED ALL ALONG THE DORSAL  
NEURAL TUBE FROM THE LEVEL OF THE DIENCEPHALON TO THE CAUDAL

CC SPINAL CORD AND THIS EXPRESSION PERSISTS UNTIL 24 HR OF  
CC DEVELOPMENT. AT THE 15-SOMITE STAGE EXPRESSION IS SEEN IN THE  
CC MIDLINE AROUND THE TAIL BUD. BETWEEN 15 AND 20 HR DEVELOPMENT  
CC DORSAL AS WELL AS VENTRAL EXPRESSION IS OBSERVED IN RECENTLY  
CC FORMED SOMITES WHILE IN MORE MATURE SOMITES, DETECTED ONLY  
CC VENTRALLY. BY 24 HR DEVELOPMENT EXPRESSION IS LIMITED TO THE  
CC VENTRAL SCLEROTOMAL ASPECT OF THE CAUDAL SOMITES. LATER IN  
CC DEVELOPMENT DETECTED IN VERY RESTRICTED PARTS OF THE CNS.  
CC -!- SIMILARITY: BELONGS TO THE NOGGIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF159147; AAD43132.1; -.  
CC ZFIN; ZDB-GENE-991206-8; nog1.  
KW Glycoprotein; Signal.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 216 NOGGIN 1.  
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 216 AA; 25093 MW; 3108242F298ABBE CRC64;  
  
Query Match 55.2%; Score 700.5; DB 1; Length 216;  
Best Local Similarity 62.6%; Pred. No. 1e-55;  
Matches 129; Conservative 28; Mismatches 36; Indels 13; Gaps 5;  
  
QY 28 QHYLHTRPAPSNLPLVDLIEHPDPIDPKEDLNETLRSLLGGHYDFGFWATSPEDR 87  
DB 21 QHYLLRPIPSDTLPLELKEDDPDIYDPRKDLNETLRSAL-GDFDSRFLSVGPPQDR 79  
QY 88 PGCGGAAGCAEDLAELDLQRLQRPSPGMPSEIKGLEFSEGLAOGKQRLSKKRLQML 147  
DB 80 YAGN-----EDLDE--QELQLNLAGMMPKDKNLDFD--APWGKKRASKKLKRLQML 128  
QY 148 WLWSOTFCPLVYAWNDLGRFWRPVYVKGVCSCFSKRSVPEGMVCKPSKSVHLTVLRWC 207  
DB 129 WLWSYFCPLVYAWNDLGRFWRPVFRAGSCYTKRSCSVPEGMVCKPAKSTHITLVRWC 188  
QY 208 -QRRGQRCGWIPYPIIIECKKSC 232  
DB 189 VARRGALKCAWIPVQYPIIIECKKSC 214  
  
RESULT 8  
ID NOGI FUGRU STANDARD; PRT; 223 AA.  
AC Q9VHT8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Noggin precursor.  
GN NOGI.  
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetraodontidae; Takifugu.  
OX NCBI\_TaxID=31033;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99102793; PubMed=9882485;  
RA Bauer H., Meier A., Hild M., Stachel S., Economides A., Hazelett D.,  
RA Harland R.M., Hammerschmidt M.;  
RT "Follistatin and noggin are excluded from the zebrafish organizer.";  
RL Dev. Biol. 204:488-507(1998).  
CC -!- FUNCTION: INHIBITOR OF BONE MORPHOGENETIC PROTEINS (BMP)  
SIGNALING.  
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Secreted.

```

CC -1- SIMILARITY: BELONGS TO THE NOGIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF095337; AAC72965.1; -.
KW Glycoprotein; Signal.
FT CHAIN 1 26 POTENTIAL.
FT CARBOHYD 27 223 NOGIN.
FT SEQUENCE 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 223 AA; 25121 MW; 35CE122E1B27F729 CRC64;

Query Match 51.7%; Score 655; DB 1; Length 223;
Best Local Similarity 53.7%; Pred. No. 1.3e-51;
Matches 123; Conservative 37; Mismatches 57; Indels 12; Gaps 5;

QY 5 PSIGVTLYALVVLGLRATPAGGQHYHIRPAPSDNLPVLDLHPDPPEKEDINET 64
DB 4 PRLRVATYLLLSVGLLHGACOPYLLLRIPEDSLPIVELKEDPQVPEDPKERDINET 63
QY 65 LRLSLGHHYDPGFMAISPPEDRRGGGGAAGADLAELDQLLRQPSGAMPEIKGLE 124
DB 64 ELKSVL-GDPDSRFLSVLPAPAE-----DGAAGN---DELDDPDAQWGGALPREIKAVD 113
QY 125 FSEGLAOGKKORLSKRLKRLQMWLMSQTFCPVLYANMDLGRFMPRYVKGSCFSKRSK 184
DB 114 F-DAPOLGKKHKSFKLRLQWLQWYSCPLAHAMTDLGRFMPRYVKGSCFSKRSK 172
QY 185 SVPEGWCKPKSKSVHLTVLRMRC-QRRGQRCGWIPIQYPIISCKKSC 232
DB 173 SVPEGWCKPKSKSVHLTVLRMRC-QRRGQRCGWIPIQYPIISCKKSC 221

RESULT 9
NOG2 BRARE STANDARD; PRT; 212 AA.
AC Q9W740;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Noggin 2 precursor.
GN NOG2.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RX MEDLINE=99423658; PubMed=10491267;
RA Fuerthner M., Thiesse B., Thiesse C.;
RA "Three different noggin genes antagonize the activity of bone
RA morphogenetic proteins in the zebrafish embryo."
RA Dev. Biol. 214:181-196(1999).
RT -1- FUNCTION: INHIBITOR OF BONE MORPHOGENETIC PROTEINS (BMP)
CC SIGNALING.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: FIRST APPEARS AT THE END OF GASTRULATION IN
CC THE AXIAL MESODERM. BY THE 5-SOMITE STAGE, EXPRESSED IN
CC TELENCEPHALON AND ANTERIOR DIENCEPHALON. FROM EARLY SEGMENTATION
CC STAGES UNTIL THE END OF TAIL ELONGATION, FOUND IN THE TAIL END.
CC EXPRESSION IS MAINTAINED DURING SCOTIOGENESIS. AT THE 10-SOMITE
CC STAGE, DETECTED IN THE ANTERO-MEDIAL ASPECT OF THE SOMITES. AT 20
CC HRS OF DEVELOPMENT EXPRESSION IS OBSERVED IN THE VENTRO-MEDIAL
CC PART OF THE SOMITE AS WELL AS IN A SMALL POPULATION OF CELLS
CC LOCATED MORE DORSALLY, ADJACENT TO THE NEURAL TUBE. AS DEVELOPMENT
CC PROCEEDS, EXPRESSION IN THE DORSAL SOMITE IS PROGRESSIVELY LOST,

```

```

CC WHILE VENTRO-MEDIAL SCLEROTOMAL CELLS LINING THE DEVELOPING AXIAL
CC VASCULATURE CONTINUE TO SHOW EXPRESSION UNTIL 30 HRS OF
CC DEVELOPMENT. IN THE HEAD, EXPRESSION IS MAINTAINED IN THE
CC TELENCEPHALON AND ANTERIOR DIENCEPHALON UNTIL LATE EMBRYOGENESIS.
CC AT THE 15-SOMITE STAGE, EXPRESSED IN THE FOREBRAIN, DORSAL
CC HINDBRAIN AND DORSAL CAUDAL SPINAL CORD. UNTIL LATE STAGES OF
CC EMBRYOGENESIS STRONG EXPRESSION IS OBSERVED IN THE DORSAL
CC HINDBRAIN WALLS.
CC -----
CC -1- SIMILARITY: BELONGS TO THE NOGIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF159148; AAD43133.1; -.
KW Glycoprotein; Signal.
FT CHAIN 1 23 POTENTIAL.
FT CARBOHYD 24 212 NOGIN 2.
FT SEQUENCE 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 212 AA; 24531 MW; D0262DD09B9D6504 CRC64;

Query Match 43.1%; Score 546; DB 1; Length 212;
Best Local Similarity 46.7%; Pred. No. 7e-42;
Matches 107; Conservative 30; Mismatches 56; Indels 36; Gaps 4;

QY 14 LVVVLGLRATPAGGQHYHIRPAPSDNLPVLDLHPDPPEKEDINETLNLSSLG 73
DB 10 LVLVLGLCHGTASQHLRLPLPSGFLVPLDINPEHPKQDSEKTLKKGSN 69
QY 74 YDPGFV-----ATSPEDRRPGGGGGAAGADLAELDQLLRQPSGAMPEIKGLE 124
DB 70 FDNFMSHLPQNLNAPAE-----LRLNPMPELKLKD 104
QY 125 FSEGLAOGKKORLSKRLKRLQMWLMSQTFCPVLYANMDLGRFMPRYVKGSCFSKRSK 184
DB 105 LTF-TPYGRVRYKVKARRKRLQWLWYTHCPVLTWKLGLRFMPRYIKENGCFSKRSK 163
QY 185 SVPEGWCKPKSKSVHLTVLRMRCQRRGQRCGWIPIQYPIISCKKSC 232
DB 164 SPPEGWCKPKSVKAVYKTFRLWYCGFMKQCTWIGVQYPIISCKKSC 212

RESULT 10
MS14 ARATH STANDARD; PRT; 507 AA.
ID MS14 ARATH
AC Q22607; Q9SLD1; Q93VF7; Q42322; Q42323;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE WD-40 repeat protein MS14.
GN MS14 OR AT2G19520 OR F3P11.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosida II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnes M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.D., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;

```



FT	ZN_FING	502	519	CCHC- <u>TYPE</u> .
FT	LIPID	2	2	MYRISTATE.
SQ	SEQUENCE	538 AA;	60858 MW; 8A7652439B464495 CRC64;	
 Query Match            7.3%; Score 92; DB 1; Length 538; Best Local Similarity 26.5%; Pred. No. 0.93; Matches      45; Conservative     12; Mismatches    45; Indels       60; Gaps          9				
OY	21 RATPACGOHY-----LHIRPAE---SDN-LPLVDLI-EHPDFDPKEDINETLLRSI	69		
Dd	122 RSPPRSSLYPALTPSLGAKPKPVQLVDSGGFLIDLTTDEPPRYDRPR-----	169		
OY	70 LGSHTPGEMATSPEPDRCGGGA--AGCAEDLAELDQLRR-----PSGAMS	118		
Dd	170 -----PPPPDRONGGEAPAGAPADPPSWASRLRRRPVPADSTTSQAFL	217		
OY	119 EIKGLEFSGLAQKKQRLSKLRLTLOWMLWQTCPVLVANNDGRF	168		
Dd	218 RAQG-----NGOLQIWPFSSS---DLYNKANNPSP	245		
 RESULT 12 MSB_MOUSE IDS MSB_MOUSE STANDARD; PRT; 1093 AA. AC Q60SI9, AD 30-MAY-2000 (Rel. 39, Created) DT 30-MAY-2000 (Rel. 39, Last sequence update) DT 16-OCT-2001 (Rel. 40, Last annotation update) DE Semaphorin 5B precursor (Semaphorin G) (Sema G). GN SEMA5B OR SEMAG OR SEMG. OS Mus musculus (Mouse). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NC NCBI_Taxid=10090; RN [1] RP SEQUENCE FROM N.A. RC STRAIN=NMT1; RX MEDLINE=96414430; PubMed=8817451; RA Adams R.H., Betz H., Pueschel A.W.; RT "A novel class of murine semaphorins with homology to thrombospondin is differentially expressed during early embryogenesis." Mech. Dev. 57:33-45(1996). RL MECH. DEV. 57:33-45(1996). CC - - FUNCTION: MAY ACT AS POSITIVE AXONAL GUIDANCE CUES. CC - - SUBCELLULAR LOCATION: Type I membrane protein. CC - - TISSUE SPECIFICITY: IN ADULT, ONLY DETECTED IN BRAIN. CC - - DEVELOPMENTAL STAGE: DIFFERENTIALLY EXPRESSED IN EMBRYONIC AND ADULT TISSUES. ITS ABUNDANCE DECREASES FROM E10 TO BIRTH. CC - - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY. CC - - SIMILARITY: CONTAINS 1 SEMA DOMAIN. CC - - SIMILARITY: CONTAINS 7 TSP TYPE-1 DOMAINS. CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL collaboration. CC The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb-sib.ch/announce/ CC or send an email to license@sib-sib.ch). CC ----- Dd EMBL; X97818; CA66398.1; -. Dd MGD; WGI:107555; Sema5b. Dr InterPro; IPRO03659; Plexin-like. Dr InterPro; IPRO02165; Plexin_repeat. Dr InterPro; IPRO01627; Sema. Dr InterPro; IPRO00884; TSP1. Dr Pfam; PF00090; TSP_1; 5. Dr Pfam; PF01403; Sema; 1. Dr Pfam; PF01437; PSI; 1. Dr SMART; SM00423; PSI; 1. Dr SMART; SM00209; TSP1; 4. Dr PROSITE; PS50092; TSP1; 6. Dr Signal; Transmembrane; Glycoprotein; Multi-pass family; Neurogenesis; KW Developmental protein; Receptor; G-protein coupled receptor.				

FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	1093	SEMAPHORIN 5B.
FT	DOMAIN	20	978	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	979	999	POTENTIAL.
FT	DOMAIN	1000	1093	CYTOSOLASMIC (POTENTIAL).
FT	DOMAIN	236	518	SEMA.
FT	DOMAIN	551	605	TSP TYPE-1 1.
FT	DOMAIN	606	663	TSP TYPE-1 2.
FT	DOMAIN	624	720	TSP TYPE-1 3.
FT	DOMAIN	721	776	TSP TYPE-1 4.
FT	DOMAIN	795	851	TSP TYPE-1 5.
FT	DOMAIN	852	908	TSP TYPE-1 6.
FT	DOMAIN	909	958	TSP TYPE-1 7.
FT	CARBOHYD	59	59	N-LINKED (GLCNAC. .)
FT	CARBOHYD	95	95	N-LINKED (GLCNAC. .)
FT	CARBOHYD	157	157	N-LINKED (GLCNAC. .)
FT	CARBOHYD	178	178	N-LINKED (GLCNAC. .)
FT	CARBOHYD	287	287	N-LINKED (GLCNAC. .)
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. .)
FT	CARBOHYD	378	378	N-LINKED (GLCNAC. .)
FT	CARBOHYD	532	532	N-LINKED (GLCNAC. .)
FT	CARBOHYD	539	539	N-LINKED (GLCNAC. .)
FT	CARBOHYD	547	547	N-LINKED (GLCNAC. .)
FT	CARBOHYD	602	602	N-LINKED (GLCNAC. .)
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. .)
FT	CARBOHYD	944	944	N-LINKED (GLCNAC. .)
SEQ	SEQUENCE	1093 AA;	120326 MW;	29B5C9B1E8108717 CRC64;
Query Match			7.0%;	Score 89; DB 1; Length 1093;
Best Local Similarity		23.8%;	Pred. No. 4;	
Matches	57;	Conservative	22;	Mismatches 78; Indels 82; Gaps 15;
QY	43	LVLDLIEHDPPIFDPEK-----KDLNETLRLSLGCHYDPGFMATSPEDRPGGGG	93	
DB	417	VVDIVAGDQDTLVHLYIGTESGLIKALS-TASSLRRCYLEE-LHLYLP-----	464	
QY	94	AAGGADLAEI-----DOLLRPPGAGPSEIKGLFSEGLAOK-----	133	
DB	465	--GRLEPRLSLILHSARALFVGLSDRYLR-----IPLERCASVHSGACLGARDYCG	516	
QY	134	---KRLSKKLRKQLQMWLMSQ--TFQVLYAMNDLGRFPRV-----KVSGCFS	180	
DB	517	WDGKQQLCSLTLEDSSNMSLMTQNTTCVNRVNRTRDGGGPGWSPMKPCBHLDGDNSSGCLC	576	
QY	181	K-RSGSVE---GMVCSFKSVHLTVLRMCORRGQR--GMPIQYPIISSECKSC	232	
DB	577	RARSCDSRRPGGLECL-GPSIHIA-----NCSNNGAMTAMSM-----ACGSI	622	
RESULT 13				
VLGB_HSVB2				
ID	VLGB_HSVB2	STANDARD;	PRT;	917 AA.
AC	P12641;			
DT	01-OCT-1989	(Rel. 12, Created)		
DT	01-AUG-1990	(Rel. 15, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Glycoprotein B-1 precursor.			
OS	Bovine herpesvirus type 2 (strain BMV) (Bovine gammaherpesvirus 1)			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Simplexvirus.			
CC	NCBI_TaxID=10296;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=88306231; PubMed=2841793;			
RA	Hammerichmidt W., Contrats F., Mankertz J., Pauli G., Ludwig H.,			
RA	Buhk H.-U.;			
RT	"Conservation of a gene cluster including glycoprotein B in bovine			
RT	herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";			
RL	Virology 165:388-405 (1988).			
RN	[2]			
RP	SEQUENCE OF 1-200 FROM N.A.			
RA	MEDLINE=88306232; PubMed=2457278;			
RA	Hammerichmidt W., Contrats F., Mankertz J., Buhk H.-U., Pauli G.,			

```

RA Ludwig H.;
RA "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418(1988).
CC -!- FUNCTION: GBL IS A 130 kDa GLYCOPROTEIN WHICH IS NECESSARY FOR THE
CC PENETRATION OF THE VIRUS INTO THE HOST CELL AND THE INDUCTION OF A
CC SYNCTIAL PHENOTYPE.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21628; AAA46053.1; -
DR EMBL; M21632; AAA46052.1; -
DR PIR; C29242; VGBEBH.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 917 GLYCOPROTEIN B-1.
FT TRANSMEM 578 594 POTENTIAL.
FT TRANSMEM 770 786 POTENTIAL.
FT TRANSMEM 795 811 POTENTIAL.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 164 164 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 505 505 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 917 AA; 101882 MW; 1B96CBF50DB4D3F3 CRC64;

Query Match 6.9%; Score 88; DB 1; Length 917;
Best Local Similarity 23.9%; Pred. No. 4;
Matches 42; Conservative 16; Mismatches 52; Indels 66; Gaps 6;

QY 38 SDNLPVLDLIEHPDPFDPKEDLNELT.-----PMEKEEPTDTPRPSRDRAPGTPKVPAMPGVTPPSGNASEPA 115
Db 61 SPSLPALDITPQP.-----PMEKEEPTDTPRPSRDRAPGTPKVPAMPGVTPPSGNASEPA 115
QY 66 -----LRLSLGLGHYDPCFMATSPEDRPGGGGAAGAEADLAELDQLLRQPSGAMP 117
Db 116 DPALRLADLRGLKSSDDNFYVCPPTG-----ATVRLLEPRPCPELP 160
QY 118 SEIKGLEFSEGLAQGKKRLSKLRKLQWLMSQTFCPVL--YAWNDLGSRWPR 171
Db 161 ---KGLNFTGEGIAVTKEKLA-PYKFKATMYKAVTVASVWSGYSYNQFMNIFEDR 212

RESULT 14
TRUN DROME
ID TRUN DROME STANDARD; PRT; 226 AA.
AC Q24155; Q9VL09;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trunk protein precursor.
GN TRK OR CG5619.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=96033803; PubMed=7590233;
RA Casanova J., Furlions M., McCormick C.A., Struhl G.;
RT "Similarities between trunk and spatze, putative extracellular
RT ligands specifying body pattern in Drosophila.";
RL Genes Dev. 9:2539-2544(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: REQUIRED FOR ACTIVITY OF THE TOR RECEPTOR, COULD BE
CC A LIGAND OF TOR. INVOLVED IN SPECIFYING TERMINAL BODY PATTERN.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U41064; AAC46953.1; -
DR EMBL; AB003628; AAF52896.1; ALT_INIT.
DR FlyBase; FBgn003751; trk.
KW Developmental protein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 226 TRUNK PROTEIN.
SQ SEQUENCE 226 AA; 26348 MW; 2CF1B5049C63521B CRC64;

Query Match 6.9%; Score 87.5; DB 1; Length 226;
Best Local Similarity 25.4%; Pred. No. 0.84;
Matches 44; Conservative 22; Mismatches 58; Indels 49; Gaps 10;

QY 60 DLNETLLRSILGLGHYDPGFMTSPEDRPGGGGAAGAEADLAELDQLLRQPSGAMP-- 117

```

```

Db      29 ELSTOGLAKTLGAFNRVMSIDPP-----GEPEKSY--HLGYRSSSYELPFY 75
Qy      118 -----SEIKLEPSE-GLAOKKOR-LSKILR-----KIQMLWS 151.
Db      76 ADSDAISVSHFPAWETNHFALVEKKEAPRSKSLRTSAFMDRVGHPRIDGFKORPW--- 132
Qy      152 QTFCEPLVAMNDLGSFRWPYVAVGSCFSCRSQVPEGMWCKSPKSVHLTVLR 204
Db      133 --ECSSKINMIDGLNFPYRISIECIA-RKCY-DHFNCKP-KSFTIKVL 180

RESULT 15
P85B_RAT
ID_P85B_RAT STANDARD; PRT; 722 AA.
AC Q63788;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
DE p85-beta subunit) (Pcdins-3-kinase p85-beta).
GN PIK3R2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96214979; PubMed=8621382;
RA Inukai K., Anai M., Vandreva E., Hosaka T., Katagiri H., Funaki M.,
RA Fukushima Y., Ogihara T., Yazaki Y., Kikuchi M., Oka Y., Amano T.,
RT "A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase
RT structurally similar to p55PIK is generated by alternative splicing
RT of the p85alpha gene."
RL J. Biol. Chem. 271:5317-5320(1996).
CC -!- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC -!- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNIT.
CC -!- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -!- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64046; BAA10926.1; -
DR HSRP: P23727; PPNB.
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000198; RhogAP.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00017; SH2; 2.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00620; RhogAP; 1.
DR PRINTS: PR00678; PI3KINSEP85.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000066; SH3; 1.
DR ProDom: PD000093; SH2; 2.
DR SMART: SM00324; RhogAP; 1.
DR SMART: SM00325; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00001; SH2; 2.
DR PROSITE: PS50002; SH3; 1.
DR SH3 domain; SH2 domain; Repeat.

```

```

FT      DOMAIN      4      80      SH3.
FT      DOMAIN      125    257    RHO-GAP.
FT      DOMAIN      324    419    SH2 1.
FT      DOMAIN      616    710    SH2 2.
SQ      SEQUENCE    722 AA; 81328 MW; 1208368B9F6FC95 CRC64;

Query Match
Best Local Similarity 25.0%; Pred. No. 3.3;
Matches 46; Conservative 22; Mismatches 63; Indels 53; Gaps 10;

Qy      3 RCPSLGTVLVALVVLG---LRATPAGQHYLHIRPAPSDNPLVLDLIEHPDIPDKK 59
Db      235 RAPSPATAVAHALASAFGPIILRAPPGKGD-GSEBAP----- 271
Qy      60 DINEETLRLSLGGHYD-----PGFMATSPEDRPG-----GGGAGAGADL----- 101
Db      272 DFPVLLIERLVQEHVDQDTAPPALPPKSKVPAFTALANGSTPPLQDAEWYGDISR 331
Qy      102 AEIDQLRQRPAGM-----PSRIKLEPSEGLAOKKORLSKILRKIQMLWSQ--TF 154
Db      332 EEWNERLRDTPDGTFLVRDASSKIQG-EYTLIRKGNKKLIKVFHRD-GHYGFSPLTF 389
Qy      155 CPVL 158
Db      390 CSVV 393

```

Search completed: January 7, 2003, 09:23:53.  
Job time : 14 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2003, 09:22:37 ; Search time 20 Seconds  
(without alignments)  
1115.159 Million cell updates/sec

Title: US-09-897-322-2  
Perfect score: 1268  
Sequence: 1 MERCPSLGVTLYALVVVLGL.....ORCGWIPQYPIISECKKSC 232

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	949.5	74.9	222	2	A43343	noggin - African c
2	96	7.6	507	2	G84577	probable WD-40 rep
3	94.5	7.5	684	2	I39595	phospholipase C -
4	92	7.3	538	1	FOVM1M	gag polyprotein -
5	88	6.9	278	2	A97471	hypothetical prote
6	88	6.9	278	2	AF2689	conserved hypothet
7	88	6.9	814	1	I39627	nicotine dehydroge
8	88	6.9	917	1	VG8EBH	glycoprotein B pre
9	87.5	6.9	728	2	H59435	phosphoinositide-3
10	86.5	6.8	457	2	G75266	UDP-N-acetylmurano
11	86.5	6.8	529	2	T45134	hypothetical prote
12	86.5	6.8	622	2	G96762	hypothetical prote
13	86	6.8	625	2	T02033	calcium/calmodulin
14	86	6.8	723	2	B38749	3-phosphatidylinos
15	86	6.8	981	1	FOVMGM	gag-abl polyprotel
16	85	6.7	1621	2	S62356	TRP-185 protein -
17	84	6.6	363	2	A72567	hypothetical prote
18	84	6.6	640	2	T08179	LRG5 protein - Chl
19	83	6.5	266	2	T29609	hypothetical prote
20	83	6.5	538	1	FOVMVM	gag polyprotein -
21	83	6.5	1066	2	G84746	hypothetical prote
22	82	6.5	122	2	C75317	hypothetical prote
23	82	6.5	344	2	E75629	probable integrase
24	82	6.5	368	2	I50233	homeobox protein -
25	82	6.5	461	2	T35151	hypothetical prote
26	82	6.5	468	1	FOVMVM	gag polyprotein -
27	81.5	6.4	436	2	H87697	mannose-1-phosphat
28	80	6.3	539	2	C81805	TspB protein NMA17
29	79.5	6.3	397	2	T17008	knottedi-like home

30	79.5	6.3	398	2	T17009	knottedi-like home
31	79.5	6.3	1310	1	W2BE62	gene 62 protein -
32	79	6.2	3511	2	A59295	unconventional myo
33	78.5	6.2	1032	2	D83637	serine/threonine p
34	78	6.2	312	2	E87350	acetyltransferase
35	78	6.2	1268	2	A49674	flightless-I homoi
36	78	6.2	1572	2	T00027	brain-specific ang
37	77.5	6.1	490	2	I41024	colicin 10 - Esche
38	77	6.1	192	2	F72805	gp49 protein - Myc
39	77	6.1	697	2	T16306	hypothetical prote
40	77	6.1	811	2	T17863	vitellogenin II ho
41	77	6.1	1234	2	S52099	phospholipase C be
42	77	6.1	1309	2	T00078	probable RNA-direc
43	76.5	6.0	230	2	T45404	hypothetical prote
44	76.5	6.0	232	2	B87122	conserved hypothet
45	76.5	6.0	1898	2	S46216	leukocyte antigen-

ALIGNMENTS

RESULT 1

A43343  
noggin - African clawed frog  
N:Alternate names: dorsaizing factor  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jul-2000  
C:Accession: A43343  
R:Smith, W.C.; Harland, R.M.  
Cell 70, 829-840, 1992  
A:Title: Expression cloning of noggin, a new dorsaizing factor localized to the Spemann  
A:Reference number: A43343; MUID:92386602; PMID:1339313  
A:Accession: A43343  
A:Molecule type: mRNA  
A:Residues: 1-222 <SMI>  
A:Cross-references: GB:M98807; NID:g214625; PIDN:AAA49916.1; PID:g214626  
A:Experimental source: embryo  
A:Note: sequence extracted from NCBI backbone (NCBIN:112834, NCBIP:112836)  
C:Superfamily: African clawed frog noggin  
C:Keywords: glycoprotein  
F:61/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	74.9%;	Score 949.5;	DB 2;	Length 222;
Best Local Similarity	78.1%;	Pred. No. 1.5e-79;		
Matches 175;	Conservative 17;	Mismatches 23;	Indels 9;	Gaps 2;
QY	9	VTLIALVVLGLRATPAGQHYLHTRPAPSDNLPLVDLIEHDPDIFDPKEKDLNETLLRS	68	
Db	8	VTIYALMVFLGRIDQGGQHYLHTRPAPSENPLVDLIEHDPDIFDPKEKDLNETLLRT	67	
QY	69	LLGGHYDPGFMTSPEDRPGGGGAAGAEPLAELDQLLRRORPSGAMPSEIKGLEFSESG	128	
Db	68	LMVGHDFPMFMTIILPEELK-----GVEDLGEULDLRLRQPSGAMPFAEIKGLEFVSG	119	
QY	129	LAQGGKQRLSKLRRKLQWLMSQTFPCPVLYAMNDLGSFRPWRYVKVVGSCFKRSCSVPE	188	
Db	120	L-QSKKRLSKLRRKLQWLMSQTFPCPVLYTNDLGTFRFWRYVKVVGSCYKRSKSCVPE	178	
QY	189	GMVCKPSSVHLTVLRWRCRRGGRCGHPIQYPIISECKKSC	232	
Db	179	GMVCKAAKSMHLTLRWRCRRVQKCAWITIQYPPVISECKKSC	222	

RESULT 2

G84577  
probable WD-40 repeat protein, MS14 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: G84577  
R:Liu, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; A.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999

```

A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <STO>
A:Cross-references: GB:AB002093; NID:g4191782; PIND:AA01051.1; GSPDB:GN00139
C:Genetics:
A:Gene: Atg19520
A:Map position: 2

  Query Match          7.6%; Score 96; DB 2; Length 507;
  Best Local Similarity 30.5%; Pred. No. 0.69;
  Matches 29; Conservative 10; Mismatches 40; Indels 16; Gaps 4;

  Qy      80  ATSPEDRPGCGGAG----GAEDLAELDQLLRPGSGMPEIKLEPSEGLAOGKKQ 135
              ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db      8  AVSPQATTPSGGTGASGPKRGRKPKTKEDS---QTPSSQQGSDVKMKE-----SGKKT 58

  Qy     136  RLKKLRKKLQMLMWSQTFPCVLYAMNDLGSRRWP 170
              ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db      59  QQSPSVDEKYSQW---KGLVPILYDMLANHNLVMP 90

RESULT 3
139595
phospholipase C - Aeromonas hydrophila
C:Species: Aeromonas hydrophila
C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 08-Oct-1999
C:Accession: I39595
R:Ingham, A.B.; Pemberton, J.M.
Curr. Microbiol. 31, 28-33, 1995
A>Title: A lipase of Aeromonas hydrophila showing nonhemolytic phospholipase C activity
A:Reference number: I39595; MUID:95284718; PMID:7767226
A:Accession: I39595
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-684 <RES>
A:Cross-references: EMBL:U14011; NID:9537628; PIND:AAA75598.1; PID:g537629

  Query Match          7.5%; Score 94.5; DB 2; Length 684;
  Best Local Similarity 25.2%; Pred. No. 1.3;
  Matches 58; Conservative 22; Mismatches 89; Indels 61; Gaps 10;

  Qy      19  GLRATPAGGQHYLIRAPSDNLPLVDLLEHPDIFDPKEDINETLLRSILGG--HYD 75
              ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db     289  GSRAVPASGNKDA-IATVAPSRPLAELRERPRFIPGRSSCLTSSPRLAFLVPGIKPRPS 347

  Qy      76  PGFMATSP-----PEDRPG---GGG-----AAGAEPLAELDQ 106
              ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db     348  PGTVPSPVRRHROCAEGVTSDRRPGRRCGSGLAWGTDGRSDSPDRAAGGEGAHRRGDP 407

  Qy     107  LLRORPGAMSEIKLEFSEGLAOGKKORLSKKLRRKLQMLMWSQTFPCVLYAMNDL-- 164
              ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db     408  HLRQGAGSR-AETRPDLPPAAGAGKATADAHLCRRRAV----CHHRCHLMSMKRLGE 462

  Qy     165  GSRFPVYVVKVSCFSKRSQSVPEGWVCKPSKSVHLTVLRMQRQRRGOR 214
              ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db     463  GERLCP-----GAGPDRCRCSKHHGEC-----RRGGDR 491

RESULT 4
FOMVIM
gag polyprotein - Moloney murine leukemia virus
N:Contains: core protein p15; core shell protein p30; inner coat protein p12; nucleoprotein
C:Species: Moloney murine leukemia virus
A>Note: host Mus sp. (mouse)
C>Date: 01-Sep-1981 #sequence_revision 27-Nov-1985 #text_change 24-Jul-1997
C:Accession: A03930
R:Shinnick, T.M.; Lerner, R.A.; Sutcliffe, J.G.
Nucleic Acids Res. 9, 543-548, 1981
A>Title: Nucleotide sequence of Moloney murine leukaemia virus.
A:Reference number: A93265; MUID:82035843; PMID:6169994

```

A:Accession: A039930  
A:Molecule type: genomic RNA  
A:Residues: 1-538 <SH>  
A:Experimental source: clone pMLV-1  
C:Comment: This protein is synthesized as a gag-pol polyprotein.  
C:Gene: gag  
C:Superfamily: mammalian retrovirus gag polyprotein I  
C:Keywords: core protein; inner coat protein; nucleoprotein; polyprotein F/2-131/Product: core protein p15 #status predicted <C15>  
F/132-215/Product: inner coat protein p12 #status predicted <C12>  
F/216-478/Product: core shell protein p30 #status predicted <C30>  
F/479-534/Product: nucleoprotein p10 #status predicted <C10>

Query Match 7.3%; Score 92; DB 1; Length 538;  
Best Local Similarity 26.5%; Pred. No. 1.7;  
Matches 45; Conservative 12; Mismatches 45; Indels 68; Gaps 9;

OY 21 RATPAGGCHY-----LHTRPAP---SDN-LPLVDLI-EHPPIRDPKRLNETLLSL 69  
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB 122 RSTPRSSLYPALTPSLGAKPKQVLSDSGGPLDILLTEDEPPYPDR----- 169  
  
OY 70 LGGHNDPGFMATSPEDRPBGSGGA--AGSADLAELTDLRLOR-----PSGAMPS 118  
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB 170 -----PPSDSDNGGEATPAGEAPEPMASRLRGREPPVADSTTSQAEPFL 217  
  
OY 119 EIKLEFESEGLAQKKQRSLSKKLRRKLMWLSQTFCVLVYANDLSRF 168  
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB 218 RAGG-----NGOLQYWPFSSS---DYVMKNNNPSF 245

RESULT 5  
A:7471  
hypothetical protein AGR\_C\_1672 [imported] - Agrobacterium tumefaciens (strain C58, Cer  
C:Species: Agrobacterium tumefaciens  
C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: A97471  
R:Goodner, B.; Hinkle, G.; Gatungu, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman  
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.  
Science 294, 2323-2328, 2001  
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
A:Reference number: A97359; PMID:11743194  
A:Accession: A97471  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-278 <KUR>  
A:Cross-references: CB:AEO07869; PIDN:AAK86722.1; PUD:g15155912; GSDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_1672  
A:Map position: circular chromosome

Query Match 6.9%; Score 88; DB 2; Length 278;  
Best Local Similarity 27.9%; Pred. No. 1.9;  
Matches 43; Conservative 19; Mismatches 74; Indels 18; Gaps 6;

OY 16 VVLGRATPAGQHLYLRPAESDNLPVLDLIEHDDPIFDPEKE-DLNETTLISLGHY 74  
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB 104 VVVEIDFKQPARRNLNTAPILTS----LSRLKERDGSAFPEDALAGERLAADPHRHGL 159  
  
OY 75 DPGFMATSPEDRPBGSGGAAGADLAELDOLLR---QRSGAMPSTIKGLE----FS 126  
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB 160 NFRVTATEPR-IASRGGEAGGADLTEAAMAARPSRAADAAGPBLSGVAIDICFE 218  
  
OY 127 EGLAQKKORL-----SKTLRRKLMWLSQTFCP 156  
|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB 219 KLETVERERLMPARSAKLLRLRALLSLRHYAP 252

RESULT 6  
AF2689  
conserved hypothetical protein Atu0918 [imported] - Agrobacterium tumefaciens (strain C  
C:Species: Agrobacterium tumefaciens  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002

C:Accession: AF2689  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Chen, G.E.; Chen, Y.; Woo, I.  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AF2689  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-278 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AAL41932.1; PID:g17739299; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu0918  
A:Map position: circular chromosome

Query Match 6.9%; Score 88; DB 2; Length 278;  
Best Local Similarity 27.9%; Pred. No. 1.9;  
Matches 43; Conservative 19; Mismatches 74; Indels 18; Gaps 6;

QY 16 VVGLRATPAGGQHYLHIRPASNPLVLDLIEHPDIPDPKEK-DLNETLLRSLGCHY 74  
DB 104 VVVEIDETQPARRNLNTAPLTS-----LSRLKERDGSAPFPEDALAAGERLAADPHRHGL 159  
QY 75 DPGFWATSPEDRPGGGGAAGAEADLADQLLR---ORPSGAMPSEIKGLE-----FS 126  
DB 160 NRVATWTNPR-IASRGKGEAGGALDLTAAAMARARFRAADMGFELSGVAIDICCFE 218  
QY 127 EGLAQKKQRL-----SKLRLRKLQMWLWSQTFCP 156  
DB 219 KGLTVERERLWPASAKULLRAALLSLARHYAP 252

RESULT 7  
139627  
nicotine dehydrogenase (EC 1.5.99.4) chain C - Arthrobacter nicotinovorans  
C:Species: Arthrobacter nicotinovorans  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 21-Jul-2000  
C:Accession: I39627; 837570  
R:Grether-Beck, S.; Igloi, G.L.; Pust, S.; Schilz, E.; Decker, K.; Brandsch, R.  
Mol. Microbiol. 13, 929-936, 1994  
A:Title: Structural analysis and molybdenum-dependent expression of the pAOI-encoded nic  
A:Reference number: I39625; MUID:95115562; PMID:7815950  
A:Accession: I39627  
A:Molecule type: DNA  
A:Residues: 1-814 <GRE>  
A:Cross-references: EMBL:X75338; NID:g665598; PIDN:CAA53088.1; PID:g406609  
C:Genetics:  
A:Gene: ndhC  
C:Superfamily: carbon-monoxide dehydrogenase large chain  
C:Keywords: molybdenum; molybdopterin; oxidoreductase  
F:226,743/Binding site: molybdopterin cytosine dinucleotide (Gln, Glu) #status predicted

Query Match 6.9%; Score 88; DB 1; Length 814;  
Best Local Similarity 22.3%; Pred. No. 6.5;  
Matches 52; Conservative 22; Mismatches 73; Indels 85; Gaps 10;

QY 34 RPAPSDNPLVDL-IEHPDIPDPKEKDLNETLLRSLGCHYDPGFWATSPEDRPGGG 92  
DB 126 RAAAEADACDLVGLLEHLEWLDPR-----KSIQGGP-----VANGERPDNVGIRG 171  
QY 93 GAAGGAEDIA-----ELQDLRQRSGAMPSEIKGLEFSEGLAQKKORLSKLRRLQ 146  
DB 172 RASFCDVDEAFSAENAHVVSALYHPGRVAAPMETRG----- 207  
QY 147 MWLWSQTFCPLVYAWNDLGRFW-----PRYVKV-----GSCFSK 181  
DB 208 -----CLADYEWTEEDRLKLVSTQMPPHYVWMLSLFLGFDSESRSEVSPDTGGGFG 259  
QY 182 RSCSVPEGMVCKPSKSVHL-TVLRWRCORRGQRCG-----WIPIQYPIISE 227

DB 260 KAHVPPEEML-MPLASKLKLTPKRVVEDRRENLLAGSHAHEQFVTIQYAANAE 311

RESULT 8  
VGBEBH  
glycoprotein B precursor - bovine herpesvirus 2 (strain BMV)  
C:Species: bovine herpesvirus 2  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 16-Jul-1999  
C:Accession: C29242  
R:Hammerschmidt, W.; Contraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.  
Virology 165, 388-405, 1988  
A:Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus t  
A:Reference number: A94381; MUID:88306231; PMID:2841793  
A:Accession: C29242  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-917 <HAM>  
A:Cross-references: GB:M21628; NID:g330752; PIDN:AAA46053.1; PID:g330753  
C:Superfamily: herpesvirus glycoprotein B  
C:Keywords: glycoprotein; transmembrane protein  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-917/Product: glycoprotein B #status predicted <GPB>  
F:578-594/Domain: transmembrane #status predicted <TM1>  
F:770-786/Domain: transmembrane #status predicted <TM2>  
F:795-811/Domain: transmembrane #status predicted <TM3>  
F:48,110,164,278,421,453,505,564,692/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 6.9%; Score 88; DB 1; Length 917;  
Best Local Similarity 23.9%; Pred. No. 7.4;  
Matches 42; Conservative 16; Mismatches 52; Indels 66; Gaps 6;

QY 38 SDNLPLVDLIEHPDIPDPKEKDLNETL-----PMEKEPDTLAPRASRDAPGTPKVPMPGVTPEPSGNASEPA 115  
DB 61 SSPLPALDLTPQ-----PMEKEPDTLAPRASRDAPGTPKVPMPGVTPEPSGNASEPA 115

QY 66 -----LRSLLGCHYDPGFWATSPEDRPGGGGAAGAEADLADQLLRQRSGAMP 117  
DB 116 DPAELRADRLGLKSGDDPNFYVCPPTG-----ATVVRLEPRFCPELP 160

QY 118 SRIKGLERSEGLAQKKORLSKLRKLQMWLWSOTFCPLV--YAWNDLGRFWPR 171  
DB 161 ---KGLNFTFEGIAVTFKENLA-PYRFPKATMYKAVTVASVWSGYSYQPMNIFEDR 212

RESULT 9  
H59435  
phosphoinositide-3-kinase regulatory beta chain [imported] - huamn  
C:Species: Homo sapiens (man)  
C:Date: 03-Jun-2002 #sequence\_revision 03-Jun-2002 #text\_change 03-Jun-2002  
C:Accession: H59435; A59436  
R:Volinia S; Patracchini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L; W  
Oncogene 7, 789-793, 1992  
A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinositol  
A:Reference number: H59435  
A:Accession: H59435  
A:Molecule type: DNA  
A:Residues: 1-728 <VOL>  
A:Cross-references: GB:NP\_005018; PID:g4826908; PIDN:NP\_005018.1  
R:Janssen, J.W.; Schleithoff, L.; Bartram, C.R.; Schulz, A.S.  
Oncogene 16, 1767-1772, 1998  
A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subunit  
A:Reference number: A59436; MUID:98241181; PMID:9582025  
A:Accession: A59436  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-728 <JAN>  
A:Cross-references: GB:NP\_005018; PID:g4826908; PIDN:NP\_005018.1

Query Match 6.9%; Score 87.5; DB 2; Length 728;  
Best Local Similarity 25.3%; Pred. No. 6.3;  
Matches 48; Conservative 23; Mismatches 60; Indels 59; Gaps 12;

QY 3 RCPGLGTVLALVVG---LRA-----TPAGGQHYLHRRPDSNLPVLVDLIEHPPI 53  
 DB 235 RAALGAVNALGATGPPLLRAPPSPSPPG-----APDGS-----BSPD 277  
 QY 54 FDPKEDLNTELLRSLGCHYD-----PGFMATSPPEDRPD-----GGGGAAGAEDL--- 101  
 DB 278 -----DPPALLVEXKLQEHLEBEQVAPRALPRPKAPAPVTLANGSGSPSLQDAEWY 331  
 QY 102 -----AELDQILRQPSGAM-----PSEIKGLEFSEGLAOGKKORLSKRLKQLQMWLW 150  
 DB 332 WGDISEEVNEKRLDTPDGTFLVRDASSKIQG-EYTLTLRKGGNNKLIXVFHRD-GHYGF 389  
 QY 151 SQ---TFCPVL 158  
 DB 390 SEPLTFCVV 399

## RESULT 10

C75266  
 UDP-N-acetylmuramoylalanine-D-glutamate ligase - Deinococcus radiodurans. (strain R1)  
 C/Species: Deinococcus radiodurans  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C/Accession: C75266  
 R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A/Reference number: A75250; MUID:20036896; PMID:10567266  
 A/Accession: C75266  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-457 <WHI>  
 A/Cross-references: GB:AE002079; GB:AE000513; NID:g6460315; PIDN:AAF12038.1; PID:g646031  
 A/Experimental source: strain R1  
 C/Genetics:  
 A/Map position: 1  
 C/Supfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 6.8%; Score 86.5; DB 2; Length 457;  
 Best Local Similarity 28.3%; Pred. No. 4.6;  
 Matches 34; Conservative 8; Mismatches 39; Indels 39; Gaps 3;

QY 9 VTLVALVVLGLRATPAGGQHYLH-----IRPAPSDNLPVLVDLIEHPPIFPKE 58  
 DB 21 VNLGGRVLLIYGLRSGRGVAHFLHGEVSAPFWHDLRAPED----- 61  
 QY 59 KDLNETLLRL-----LGHYDGFMAATSPPEDRPGGGGAAGAEADLAEIDLQRLR 112  
 DB 62 ---EALMRQLGHRQADLGGTYDLVVAAPGVPIIDHRLRYLAGKAEIIGEVALLARLP 117

## RESULT 11

T45134  
 hypothetical protein [imported] - Microbacterium ammoniaphilum (fragment)  
 C/Species: Microbacterium ammoniaphilum  
 C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
 C/Accession: T45134  
 R.Striebel, H.M.; Seebert, S.; Jarsch, M.; Kessler, C.  
 Gene 172, 41-46, 1996  
 A>Title: Cloning and characterization of the Mami restriction-modification system from M  
 A/Reference number: Z22923; MUID:196257250; PMID:8654988  
 A/Accession: T45134  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-529 <STR>  
 A/Cross-references: EMBL:X79027; NID:g984667; PIDN:CAA55649.1; PID:g1679831  
 A/Experimental source: ATCC 15354

Query Match 6.8%; Score 86.5; DB 2; Length 529;  
 Best Local Similarity 23.9%; Pred. No. 5.4;

Matches 47; Conservative 18; Mismatches 73; Indels 59; Gaps 8;  
 QY 19 GLATPAGGQHYLHRRP--PSNLPVLVDLIEHPDIFDPKEDLNTELLRSLGCHYD 76  
 DB 321 GVGARPARPRRLHPRADLPHGGQGVARLDHPQLEGAAGSRPRHRLRL--HQGR 378  
 QY 77 GFMAATSPEDR-----PGGGGAAGAEDLAEIDLQRLR--RPSGAMPSEIKGLEFSEGL 129  
 DB 379 GDLIRRRRDLRRGRPRGQGAHGG-----QGLRRAGRRRGVPSQLAVLR--QGV 428  
 QY 130 AGGKKQRLSKLRLKQLQMWLMSQTFCEPVLVAMNDLGSREMPRYVKGSCFSKSCSVPEG 189  
 DB 429 CQAS-----ATLALMTSGRLREL-----AG 450  
 QY 190 MCKPSKSVHLTVLRWR 206  
 DB 451 LLCRDVSFSLDVTWR 467

## RESULT 12

G96762  
 hypothetical protein FED5.2 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C/Accession: G96762  
 R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
 ansen, N.P.; Hughes, B.; Hultzer, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maltli, R.; Marziani  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.M.  
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: G96762  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-622 <STO>  
 A/Cross-references: GB:AE005173; NID:g10092369; PIDN:AA012777.1; GSPDB:GN00141  
 C/Genetics:  
 A/Map position: 1

Query Match 6.8%; Score 86.5; DB 2; Length 622;  
 Best Local Similarity 23.2%; Pred. No. 6.5;  
 Matches 48; Conservative 21; Mismatches 85; Indels 53; Gaps 9;

QY 19 GLATPAGGQHYLHRRPDSNLPVLVDLIEHPDIFDPKEDLNTELLRSLL---CG--- 72  
 DB 223 GLSATP-----RPSNLTNAETYSLOSSRNP--TPRGSSFNHTDYSWMAAGSGGNS 271  
 QY 73 HYDPG--FMATSPEDRP-----GGGGAAGAEDLAE 103  
 DB 272 NFGPGAVFGSKPTPRPNYBEEDEGPAKPTAAGTAGAGRHVYSGSGGGGGAHYPAP 331  
 QY 104 LDQLLRQPSGAMPSEIKGLEFSEGLAOGKKORLSKRLKQLQMWLMSQTFCEPVLVAMND 163  
 DB 332 NPMFSPNTGGGGGTAAK---NAPVGGKRQDGN---GRDLHFMVWSSASPVSDVGVG 385  
 QY 164 LGSRFWRPRYVKGSCFSKR--SCSVPEG 189  
 DB 386 GGGNHADYSTATNDHQKQKISVPG 412

## RESULT 13

T02033  
 calcium/calmodulin-dependent protein kinase homolog - maize  
 C/Species: Zea mays (maize)  
 C/Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 17-Nov-2000  
 C/Accession: T02033; T02994; T01694  
 R.Lu, Y.T.; Hidaka, H.; Feldman, L.J.

Search completed: January 7, 2003, 09:25:01  
Job time : 24 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 7, 2003, 09:19:22 ; Search time 37 Seconds  
(without alignments)  
835.517 Million cell updates/sec

Title: US-09-897-322-2

Perfect score: 1268

Sequence: 1 MERCPSLGVTLYALVVLGL.....QRCGWIPQYPIISECKKC 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A. Geneseq 101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1268	100.0	232	20 AAW96279	Spemann organiser
2	1268	100.0	232	23 AAG79341	Human Noggin. Hom
3	1261	99.4	232	15 AAR50303	Human noggin. Hom
4	1258	99.2	232	15 AAR50304	Murine noggin. Mu
5	1258	99.2	232	20 AAW96280	Spemann organiser
6	1258	99.2	232	23 AAG79347	Partial mouse Nogg
7	1195	94.2	449	20 AAW96278	Human noggin/immun
8	112	8.8	361	21 AAB32696	Eucalyptus grandis
9	102	8.0	2392	21 AAB07665	Amino acid sequenc
10	99.5	7.8	413	21 AAY53820	Amino acid sequenc

11	99	7.8	836	22	AA85703	Novel protein kina
12	93.5	7.4	372	18	AAW22065	Chick fringe A (ra
13	92.5	7.3	319	18	AAW18658	Fragmented human N
14	92	7.3	538	21	AAW10043	MMLV gag protein.
15	92	7.3	648	20	AAW17946	MMLV gag gene pro
16	92	7.3	761	22	ABG02677	Novel human diagno
17	92	7.3	1737	21	AAW10044	MMLV gag-pol prote
18	91	7.2	743	22	AAW79738	Human protein SEQ
19	90.5	7.1	623	22	AAW85583	Rice CDPK (clone r
20	89	7.0	1093	22	AAE02455	Mouse semaphorin G
21	88.5	7.0	522	19	AAW69341	Secreted protein o
22	88.5	7.0	623	17	AAW03140	Human IgG CHI-hing
23	88.5	7.0	979	19	AAW40309	Human ITAK protein
24	88.5	7.0	979	21	AAW95294	Human GEF containi
25	87.5	6.9	235	22	ABB61683	Drosophila melanog
26	87.5	6.9	728	22	AAW78754	Human protein SEQ
27	87.5	6.9	1368	23	AAU77182	Human kinesin moto
28	87	6.9	14	15	AAW49838	Noggin conserved p
29	87	6.9	14	15	AAW50305	Noggin conserved r
30	87	6.9	14	20	AAW96281	Conserved sequence
31	87	6.9	14	23	AAG79342	Noggin conserved p
32	87	6.9	1975	22	ABB62094	Drosophila melanog
33	86.5	6.8	622	21	AAW26950	Auxin transport pr
34	86.5	6.8	622	23	ABB91576	Herbicidally activ
35	86.5	6.8	939	23	AAG68296	Human semaphorin G
36	86.5	6.8	954	23	AAG68295	Human semaphorin G
37	86.5	6.8	999	21	AAW94990	Human secreted pro
38	86.5	6.8	1034	23	AAG68291	Human semaphorin G
39	86.5	6.8	1049	23	AAG68289	Human semaphorin G
40	86.5	6.8	1078	23	AAG68292	Human semaphorin G
41	86.5	6.8	1092	23	ABG34077	Human Pro peptide
42	86.5	6.8	1093	23	AAG68290	Human semaphorin G
43	86.5	6.8	1136	23	AAG68294	Human semaphorin G
44	86.5	6.8	1151	23	AAG68293	Human semaphorin G
45	86.5	6.8	1352	23	AAE19212	Human MOL4 protein

ALIGNMENTS

RESULT 1

AAW96279

ID AAW96279 standard; Protein; 232 AA.

XX AAW96279;

XX 14-JUN-1999 (first entry)

XX Spemann organiser signal Noggin polypeptide.

XX Spemann organiser signal; noggin; bone morphogenetic protein; BMP;  
KW disease; disorder; bone; bone growth; trauma; burns;  
KW Fibrodysplasia Ossificans Progressiva; FOP; antibodies; detection;  
KW diagnosis; therapy.

XX Homo sapiens.

XX WO9503996-A1.

XX 28-JAN-1999.

XX 17-JUL-1998; 98WO-US14603.

XX 17-JUL-1997; 97US-0897236.

XX (REGE-) REGENERON PHARM INC.

XX (REGC ) UNIV CALIFORNIA.

XX Economides AN, Harland RM, Stahl N;

XX WPI; 1999-132240/11.

XX N-PSDB; AAX09017.

[illegible]

XX	(REGE ) REGENERON PHARM INC.
PA	(RECG ) UNIV CALIFORNIA.
PI	Valenzuela DM, Ip NY, Cudny HD, Yancopoulos GD, Harland RM,
PI	Smith WC, Lamb T, Knecht A;
XX	
XX	MPI; 2002-442065/47.
DR	N-P8DB; AA172927.
XX	
PS	Recombinant preparation of noggin polypeptide, useful e.g. for treating neurodegenerative disease, by expressing specific human sequence -
PT	
XX	
PS	Example 4; Fig 1; 40pp; English.
XX	
CC	This sequence represents human Noggin polypeptide. Noggin is a
CC	neurotrophic growth factor which induces dorsal development in vertebrates. Noggin modifies the actions of fibroblast growth factor (FGF). It is a 26 kD protein with a hydrophobic amino terminal. Noggin is secreted, apparently as a dimeric glycoprotein. The carboxy terminal region of Noggin shows homology to a kunitz-type protease inhibitor.
CC	Noggin polypeptide may be prepared by culturing cells transformed with a vector that contains a control sequence operatively linked to a nucleic acid molecule which comprises the coding region for human noggin or a sequence encoding the same amino acid sequence. Human
CC	Noggin, also its fusion proteins and derivatives, may be used to raise specific antibodies (Ab), for diagnosis, for detection and purification of Ab, to induce growth of nerve and muscle cells in mammals, and to regulate bone or muscle growth, e.g. in wound-healing compositions and for treating neurodegeneration (Alzheimer's, Parkinson's or Huntington's diseases, amyotrophic lateral sclerosis and peripheral neuropathy), traumatic nerve injury, diabetes, kidney dysfunction, the toxic effects of chemotherapeutic agents being used to treat acquired immune deficiency syndrome or cancer, and congenital malformations such as anencephaly, as an additive to culture media used for growing nerve cells and to isolate cognate receptors, potentially useful for diagnosis of some cancers. Abs are used for in vitro or in vivo therapy or diagnosis and for purification of Noggin.
CC	
XX	
SQ	Sequence 232 AA;
	Query Match: 100.0%; Score 1268; DB 23; Length 232;
	Best Local Similarity 100.0%; Pred. NO. 5e-123;
	Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 MERCPSLGVTTLVAVVGLRATPAGGOHYLHIRAPSDNLEPLVDLIIEHPDIIDPKKED 60
Db	1 MERCPSLGVTTLVAVVGLRATPAGGOHYLHIRAPSDNLEPLVDLIIEHPDIIDPKKED 60
OY	61 INETLLRLSLGGHYDPGFMAISPEPDPRGGGGGAAGAEDLAELDQLRORSGAMPSEI 120
Db	61 INETLLRLSLGGHYDPGFMAISPEPDPRGGGGGAAGAEDLAELDQLRORSGAMPSEI 120
OY	121 KGLESEGAGOGKKORLSKLRRKIOMWMSOTPCPVLYAMWDLSRFMPRVVXGSCFS 180
Db	121 KGLESEGAGOGKKORLSKLRRKIOMWMSOTPCPVLYAMWDLSRFMPRVVXGSCFS 180
OY	181 KRSCVPESGMCKPSKSVAHLTVLRMRCQRGGQRGMVIPQYPIISECKCSC 232
Db	181 KRSCVPESGMCKPSKSVAHLTVLRMRCQRGGQRGMVIPQYPIISECKCSC 232
RESULT 3	
ID	AAR50303
AA	AAR50303 standard; Protein; 232 AA.
XX	
AC	AAR50303;
XX	
DT	19-OCT-1994 (first entry)
XX	
DE	Human noggin.
XX	
KW	Human; noggin; hydrophobic amino terminal; kunitz-type; bone growth;

KW protease inhibitor; regulation; cartilage; growth factor; epidermis;  
KW tissue matrix; potentiation; wound healing; diagnosis; probe; tumour;  
KW fibroblast growth factor; FGF; activin; nerve; muscle cell;  
KW Alzheimers disease; Parkinsons disease; Huntington's chorea;  
KW peripheral neuropathy.

XX Homo sapiens.

OS W09405791-A.

XX 17-MAR-1994.

XX 02-SEP-1993; 93WO-US08326.

XX 03-SEP-1992; 92US-0939954.

XX 23-SEP-1992; 92US-0950410.

XX 06-OCT-1992; 92US-0957401.

XX (REGE-) REGENERON PHARM INC.

XX Cudny H, Harland RM, Ip NY, Knecht A, Lamb T, Smith WC;

XX Valenzuela DM, Yancopoulos GD;

XX WPI: 1994-101196/12.

XX N-PSDB; AAQ76342.

XX Noggin protein capable of inducing dorsal growth, and sequences

XX encoding it - useful for treating neurodegenerative disorders and

XX neural damage, e.g. due to trauma or after chemotherapy

XX Claim 1; Fig 1; 100pp; English.

XX This sequence represents human noggin protein. The noggin cDNA

XX encodes a 26 kD secreted protein which has a hydrophobic amino

XX terminal sequence. The carboxy terminal sequence of noggin shows

XX homology to a kunitz-type protease inhibitor, indicating that it may

XX exhibit activities of a protease inhibitor. Noggin is a regulator of

XX cartilage production and a growth factor for tissue matrix and

XX epidermis. Noggin is useful for regulating cartilage and bone growth,

XX optionally in conjunction with other growth factors which may be

XX potentiated by noggin. It is also useful in wound healing and in the

XX isolation of its receptor, which may itself be used as a diagnostic

XX probe for certain types of tumour. Noggin modifies the actions of

XX fibroblast growth factor (FGF) and also activin. Noggin may be used

XX for enhancing the survival or inducing the growth of nerve and muscle

XX cells. It may therefore be useful in the therapy of congenital

XX conditions or degenerative disorders of the nervous system, eg.

XX Alzheimers disease, Parkinsons disease, Huntington's chorea and or

XX peripheral neuropathy.

XX Sequence 232 AA;

XX Query Match 99.4%; Score 1261; DB 15; Length 232;

XX Best Local Similarity 99.6%; Pred. No. 2.6e-122;

XX Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYLRAPSDNPLVDLIEHPDFDPKXD 60

XX 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYLRAPSDNPLVDLIEHPDFDPKXD 60

XX 61 LNETLRLSLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

XX Homo sapiens.

OS W09405791-A.

XX 17-MAR-1994.

XX 02-SEP-1993; 93WO-US08326.

XX 03-SEP-1992; 92US-0939954.

XX 23-SEP-1992; 92US-0950410.

XX 06-OCT-1992; 92US-0957401.

XX (REGE-) REGENERON PHARM INC.

XX Cudny H, Harland RM, Ip NY, Knecht A, Lamb T, Smith WC;

XX Valenzuela DM, Yancopoulos GD;

XX WPI: 1994-101196/12.

XX N-PSDB; AAQ76342.

XX Noggin protein capable of inducing dorsal growth, and sequences

XX encoding it - useful for treating neurodegenerative disorders and

XX neural damage, e.g. due to trauma or after chemotherapy

XX Claim 1; Fig 1; 100pp; English.

XX This sequence represents human noggin protein. The noggin cDNA

XX encodes a 26 kD secreted protein which has a hydrophobic amino

XX terminal sequence. The carboxy terminal sequence of noggin shows

XX homology to a kunitz-type protease inhibitor, indicating that it may

XX exhibit activities of a protease inhibitor. Noggin is a regulator of

XX cartilage production and a growth factor for tissue matrix and

XX epidermis. Noggin is useful for regulating cartilage and bone growth,

XX optionally in conjunction with other growth factors which may be

XX potentiated by noggin. It is also useful in wound healing and in the

XX isolation of its receptor, which may itself be used as a diagnostic

XX probe for certain types of tumour. Noggin modifies the actions of

XX fibroblast growth factor (FGF) and also activin. Noggin may be used

XX for enhancing the survival or inducing the growth of nerve and muscle

XX cells. It may therefore be useful in the therapy of congenital

XX conditions or degenerative disorders of the nervous system, eg.

XX Alzheimers disease, Parkinsons disease, Huntington's chorea and or

XX peripheral neuropathy.

XX Sequence 232 AA;

XX Query Match 99.4%; Score 1261; DB 15; Length 232;

XX Best Local Similarity 99.6%; Pred. No. 2.6e-122;

XX Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYLRAPSDNPLVDLIEHPDFDPKXD 60

XX 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYLRAPSDNPLVDLIEHPDFDPKXD 60

XX 61 LNETLRLSLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

XX 61 LNETLRLSLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

XX 121 KGLEFSEGLAOGKQRLSKLRLKQWLWNSQTFPCVLYAWNDLGRFRPYYKVGSCFS 180

XX 121 KGLEFSEGLAOGKQRLSKLRLKQWLWNSQTFPCVLYAWNDLGRFRPYYKVGSCFS 180

XX Homo sapiens.

OS W09405791-A.

XX 17-MAR-1994.

XX 02-SEP-1993; 93WO-US08326.

XX 03-SEP-1992; 92US-0939954.

XX 23-SEP-1992; 92US-0950410.

XX 06-OCT-1992; 92US-0957401.

XX (REGE-) REGENERON PHARM INC.

XX Cudny H, Harland RM, Ip NY, Knecht A, Lamb T, Smith WC;

XX Valenzuela DM, Yancopoulos GD;

XX WPI: 1994-101196/12.

XX N-PSDB; AAQ76342.

XX Noggin protein capable of inducing dorsal growth, and sequences

XX encoding it - useful for treating neurodegenerative disorders and

XX neural damage, e.g. due to trauma or after chemotherapy

XX Claim 1; Fig 1; 100pp; English.

XX This sequence represents human noggin protein. The noggin cDNA

XX encodes a 26 kD secreted protein which has a hydrophobic amino

XX terminal sequence. The carboxy terminal sequence of noggin shows

XX homology to a kunitz-type protease inhibitor, indicating that it may

XX exhibit activities of a protease inhibitor. Noggin is a regulator of

XX cartilage production and a growth factor for tissue matrix and

XX epidermis. Noggin is useful for regulating cartilage and bone growth,

XX optionally in conjunction with other growth factors which may be

XX potentiated by noggin. It is also useful in wound healing and in the

XX isolation of its receptor, which may itself be used as a diagnostic

XX probe for certain types of tumour. Noggin modifies the actions of

XX fibroblast growth factor (FGF) and also activin. Noggin may be used

XX for enhancing the survival or inducing the growth of nerve and muscle

XX cells. It may therefore be useful in the therapy of congenital

XX conditions or degenerative disorders of the nervous system, eg.

XX Alzheimers disease, Parkinsons disease, Huntington's chorea and or

XX peripheral neuropathy.

XX Sequence 232 AA;

XX Query Match 99.4%; Score 1261; DB 15; Length 232;

XX Best Local Similarity 99.6%; Pred. No. 2.6e-122;

XX Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYLRAPSDNPLVDLIEHPDFDPKXD 60

XX 1 MERCPSLGVTLYALVVVLGLRATPAGQGHYLRAPSDNPLVDLIEHPDFDPKXD 60

XX 61 LNETLRLSLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

XX 61 LNETLRLSLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120

XX 121 KGLEFSEGLAOGKQRLSKLRLKQWLWNSQTFPCVLYAWNDLGRFRPYYKVGSCFS 180

XX 121 KGLEFSEGLAOGKQRLSKLRLKQWLWNSQTFPCVLYAWNDLGRFRPYYKVGSCFS 180

Db 61 INETLLRSLILGHYDPEFMATSPEDRPGGGGAGAGADLAELDQLLRQSPGAMPSSEI 120  
 Qy 121 KGLEFSEBGLAOGKKQRLSKTLRRKLQMWLMSQTFCEVLYAMNDLGSRFMPRYVYVSGCFS 180  
 Db 121 KGLEFSEBGLAOGKKQRLSKTLRRKLQMWLMSQTFCEVLYAMNDLGSRFMPRYVYVSGCFS 180  
 Qy 181 KRCSVPGEWVCKPSKSVHLTVLRMRCORRGRCGMIPIQYPIISECKSC 232  
 Db 181 KRCSVPGEWVCKPSKSVHLTVLRMRCORRGRCGMIPIQYPIISECKSC 232  
 RESULT 5  
 AA96280  
 ID AA96280 standard; protein; 232 AA.  
 AC AA96280;  
 DT 14-JUN-1999 (first entry)  
 DE Spemann organiser signal Noggin polypeptide.  
 XX Spemann organiser signal; noggin; bone morphogenetic protein; BMP;  
 KW disease; disorder; bone; bone growth; trauma; burns;  
 KW Fibrodysplasia Ossificans Progressiva; FOP; antibodies; detection;  
 KW diagnosis; therapy.  
 XX Mus musculus.  
 OS Mus musculus.  
 PN WO903996-A1.  
 PD 28-JAN-1999.  
 PF 17-JUL-1998; 98WO-US14603.  
 PR 17-JUL-1997; 97US-0897236.  
 PA (REGC-) REGENERON PHARM INC.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Economides AN, Harland RM, Stahl N;  
 DR WPI; 1999-132240/11.  
 DR N-PSDB; AAX09018.  
 PT New modified Noggin polypeptides - having an amino acid deletion to  
 PT increase bioavailability in animal serum, used for treating bone  
 PT disorders or diseases  
 PS Disclosure; Figure 13; 132pp; English.  
 XX Disclosures; Figure 13; 132pp; English.  
 CC Noggin polypeptides induce dorsal growth and block bone  
 CC morphogenetic protein (BMP) activity in vertebrates. Modified Noggin  
 CC polypeptides, modified by a deletion of amino acids 138-144, can be  
 CC used for the treatment of a disease or disorder affecting bone.  
 CC e.g. abnormal bone growth following hip replacement surgery, trauma,  
 CC burns, or spinal cord injury, or Fibrodysplasia Ossificans  
 CC Progressiva (FOP). This is achieved by the modified Noggin  
 CC polypeptide exhibiting improved bioavailability in animal  
 CC sera while retaining the ability to bind to a BMP. The products can  
 CC also be used for the production of antibodies, detection and  
 CC diagnosis. The modified Noggin protein has enhanced therapeutic  
 CC properties.  
 XX Sequence 232 AA;  
 SQ  
 Query Match 99.2%; Score 1258; DB 20; Length 232;  
 Best Local Similarity 99.1%; Pred. No. 5, 4e-122;  
 Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 61 INETLLRSLILGHYDPEFMATSPEDRPGGGGAGAGADLAELDQLLRQSPGAMPSSEI 120  
 Db 61 INETLLRSLILGHYDPEFMATSPEDRPGGGGAGAGADLAELDQLLRQSPGAMPSSEI 120  
 Qy 121 KGLEFSEBGLAOGKKQRLSKTLRRKLQMWLMSQTFCEVLYAMNDLGSRFMPRYVYVSGCFS 180  
 Db 121 KGLEFSEBGLAOGKKQRLSKTLRRKLQMWLMSQTFCEVLYAMNDLGSRFMPRYVYVSGCFS 180  
 Qy 181 KRCSVPGEWVCKPSKSVHLTVLRMRCORRGRCGMIPIQYPIISECKSC 232  
 Db 181 KRCSVPGEWVCKPSKSVHLTVLRMRCORRGRCGMIPIQYPIISECKSC 232  
 RESULT 6  
 AAG79347  
 ID AAG79347 standard; protein; 232 AA.  
 AC AAG79347;  
 DT 21-AUG-2002 (first entry)  
 DE Partial mouse Noggin.  
 XX Human; noggin; neurotrophic; growth factor; dorsal development;  
 KW vertebrate; fibroblast growth factor; FGF; cognate receptor; cancer;  
 KW Kunitz-type protease inhibitor; nerve; muscle; bone; neurodegeneration;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;  
 KW amyotrophic lateral sclerosis; peripheral neuropathy; culture media;  
 KW traumatic nerve injury; diabetes; kidney dysfunction; anencephaly.  
 XX Mus musculus.  
 OS Mus musculus.  
 PN US6277593-B1.  
 PD 21-AUG-2001.  
 PF 07-OCT-1998; 98US-0167874.  
 PR 02-SEP-1993; 93US-0392935.  
 PR 02-SEP-1993; 93WO-US08326.  
 PR 07-JUN-1995; 95US-0485721.  
 PR 03-SEP-1992; 92US-0939954.  
 PR 23-SEP-1992; 92US-0950410.  
 PR 06-OCT-1992; 92US-0957401.  
 PA (REGC-) REGENERON PHARM INC.  
 PA (REGC) UNIV CALIFORNIA.  
 PI Valenzuela DM, Ip NY, Cuddy HD, Yancopoulos GD, Harland RM;  
 PI Smith WC, Lamb T, Knecht A;  
 DR WPI; 2002-442065/47.  
 DR N-PSDB; AAI72930.  
 PT Recombinant preparation of noggin polypeptide, useful e.g. for treating  
 PT neurodegenerative disease, by expressing specific human sequence -  
 PS Example 3; Fig 13; 40pp; English.  
 XX This sequence represents partial mouse Noggin polypeptide. Noggin is  
 CC a neurotrophic growth factor which induces dorsal development in  
 CC vertebrates. Noggin modifies the actions of fibroblast growth factor  
 CC (FGF). It is a 26 kD protein with a hydrophobic amino terminal. Noggin  
 CC is secreted, apparently as a dimeric glycoprotein. The carboxy terminal  
 CC region of Noggin shows homology to a kunitz-type protease inhibitor.  
 CC Noggin polypeptide may be prepared by culturing cells transformed with  
 CC a vector that contains a control sequence operatively linked to a  
 CC nucleic acid molecule which comprises the coding region for human  
 CC noggin or a sequence encoding the same amino acid sequence. Human  
 CC Noggin, also its fusion proteins and derivatives, may be used to raise  
 CC specific antibodies (Ab), for diagnosis, for detection and purification  
 CC of Ab, to induce growth of nerve and muscle cells in mammals, and to  
 CC regulate bone or muscle growth, e.g. in wound-healing compositions and

CC for treating neurodegeneration (Alzheimer's, Parkinson's or  
 CC Huntington's diseases, amyotrophic lateral sclerosis and peripheral  
 CC neuropathy), traumatic nerve injury, diabetes, kidney dysfunction, the  
 CC toxic effects of chemotherapeutic agents being used to treat acquired  
 CC immune deficiency syndrome or cancer, and congenital malformations such  
 CC as anencephaly, as an additive to culture media used for growing nerve  
 CC cells and to isolate cognate receptors, potentially useful for  
 CC diagnosis of some cancers. Ab's are used for in vitro or in vivo  
 CC therapy or diagnosis and for purification of Noggin.  
 XX  
 SQ Sequence 232 AA;

Query Match 99.2%; Score 1258; DB 23; Length 232;  
 Best Local Similarity 99.1%; Pred. No. 5.4e-122;  
 Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MERCPSLGVTLYALVVGLRATPAGGQHYLHIRPAPSDNPLVDLIEHPDPIFDPEKD 60  
 |||||  
 Db 1 MERCPSLGVTLYALVVGLRATPAGGQHYLHIRPAPSDNPLVDLIEHPDPIFDPEKD 60  
 |||||  
 QY 61 LNETLLRSLLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120  
 |||||  
 Db 61 LNETLLRSLLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120  
 |||||  
 QY 121 KGLEFSEGLAQKKQRLSKLRKQLQMWLWSQTFCPVLYAWNDLGSRFWRPVRVYKVGSCFS 180  
 |||||  
 Db 121 KGLEFSEGLAQKKQRLSKLRKQLQMWLWSQTFCPVLYAWNDLGSRFWRPVRVYKVGSCFS 180  
 |||||  
 QY 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCRRGQRCGWIPQIPIIIECKKSC 232  
 |||||  
 Db 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCRRGQRCGWIPQIPIIIECKKSC 232  
 |||||

RESULT 7  
 AAW96278  
 ID AAW96278 standard; Protein; 449 AA.

XX AAW96278;  
 XX  
 XX  
 DT 14-JUN-1999 (first entry)  
 DE Human noggin/immunoglobulin G1 fusion peptide (hNGdelta133-144Fc).  
 DE  
 KW Spemann organizer signal; noggin; bone morphogenetic protein; BMP;  
 KW disease; disorder; bone; bone growth; trauma; burns;  
 KW Fibrodysplasia Ossificans Progressiva; FOP; antibodies; detection;  
 KW diagnosis; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..30  
 FT /label= Putative\_signal\_peptide  
 FT 62  
 FT Modified-site /label= Glycosylation\_site  
 FT 133..134  
 FT /note= "Marks position of 133-144 deletion"  
 FT Disulfide-bond 218  
 FT /note= "Cysteine residue used in inter-chain  
 FT disulfide bridge in human noggin"  
 FT 221..222  
 FT /note= "Ser-Gly bridge connecting human noggin  
 FT sequence to human IgG1 Fc region"  
 FT Disulfide-bond 228  
 FT /note= "Cysteine involved in inter-chain  
 FT disulfide bridge of IgG hinge preceding  
 FT the Fc region"  
 FT Disulfide-bond 232  
 FT /note= "Cysteine involved in inter-chain  
 FT disulfide bridge of IgG hinge preceding  
 FT the Fc region"  
 FT 223..449  
 FT /label= Fc\_domain

FT Modified-site 299  
 FT /label= Glycosylation\_site  
 PN WO9903996-A1.  
 XX  
 PD 28-JAN-1999.  
 XX  
 PF 17-JUL-1998; 98WO-US14603.  
 XX  
 PR 17-JUL-1997; 97US-0897236.  
 XX  
 PA (REGC-) REGENERON PHARM INC.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Economides AN, Harland RM, Stahl N;  
 XX  
 DR WPI; 1999-132240/11.  
 XX  
 PT New modified Noggin polypeptides - having an amino acid deletion to  
 PT increase bioavailability in animal serum, used for treating bone  
 PT disorders or diseases  
 XX  
 PS Example 9; Figure 14; 132pp; English.  
 XX  
 CC Noggin polypeptides induce dorsal growth and block bone  
 CC morphogenetic protein (BMP) activity in vertebrates. Modified Noggin  
 CC polypeptides, modified by a deletion of amino acids 138-144, can be  
 CC used for the treatment of a disease or disorder effecting bone.  
 CC e.g. abnormal bone growth following hip replacement surgery, trauma,  
 CC burns, or spinal cord injury, or Fibrodysplasia Ossificans  
 CC Progressiva (FOP). This is achieved by the modified Noggin  
 CC polypeptide exhibiting improved bioavailability in animal  
 CC sera while retaining the ability to bind to a BMP. The products can  
 CC also be used for the production of antibodies. detection and  
 CC diagnosis. The modified Noggin protein (AAW96278) has enhanced  
 CC therapeutic properties.  
 XX  
 SQ Sequence 449 AA;

Query Match 94.2%; Score 1195; DB 20; Length 449;  
 Best Local Similarity 94.8%; Pred. No. 4.5e-115;  
 Matches 220; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

QY 1 MERCPSLGVTLYALVVGLRATPAGGQHYLHIRPAPSDNPLVDLIEHPDPIFDPEKD 60  
 |||||  
 Db 1 MERCPSLGVTLYALVVGLRATPAGGQHYLHIRPAPSDNPLVDLIEHPDPIFDPEKD 60  
 |||||  
 QY 61 LNETLLRSLLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120  
 |||||  
 Db 61 LNETLLRSLLGGHYDPGFMTATPPEDRPGGGGAAGGAEDLAELDQLLRQPSGAMPSEI 120  
 |||||  
 QY 121 KGLEFSEGLAQKKQRLSKLRKQLQMWLWSQTFCPVLYAWNDLGSRFWRPVRVYKVGSCFS 180  
 |||||  
 Db 121 KGLEFSEGLAQK-----LQMWLWSQTFCPVLYAWNDLGSRFWRPVRVYKVGSCFS 168  
 |||||  
 QY 181 KRSCSVPEGMVCKPSKSVHLTVLRWRCRRGQRCGWIPQIPIIIECKKSC 232  
 |||||  
 Db 169 KRSCSVPEGMVCKPSKSVHLTVLRWRCRRGQRCGWIPQIPIIIECKKSC 220  
 |||||

RESULT 8  
 AAB32696  
 ID AAB32696 standard; Protein; 361 AA.  
 XX  
 AC AAB32696;  
 XX  
 DT 25-JAN-2001 (first entry)  
 XX  
 DE Eucalyptus grandis transcription factor protein sequence #154.  
 XX  
 KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
 KW poplar; sweetgum; teak; mahogany; bz1p; G-box binding factor;  
 KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;

KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;  
 KW type 2 Cys2His2; CCAAT box element; MYB.  
 OS Eucalyptus grandis.  
 PN WO20005724-A2.  
 PD 14-SEP-2000.  
 XX 09-MAR-2000; 2000WO-US06112.  
 XX 11-MAR-1999; 99US-0266513.  
 PR 18-AUG-1999; 99US-0149485.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLEET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX Wood M, McGrath A, Shenk MA, Glenn M;  
 PI WPI; 2000-579369/54.  
 DR New isolated polynucleotide encoding a plant transcription factor for  
 XX producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT -  
 XX  
 XX Claim 8; Page 261; 747bp; English.  
 PS The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such a  
 CC transcription factor. The transcription factor may be used to produce a  
 CC plant having modified gene expression such as a woody plant e.g. a  
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
 CC to modify the activity of a polypeptide in a plant. The transcription  
 CC factors of the present invention are members from the following families  
 CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic  
 CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain  
 CC zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2  
 CC Cys2His2, CCAAT box elements and MYB.  
 XX  
 SQ Sequence 361 AA;  
 Query Match 8.8%; Score 112; DB 21; Length 361;  
 Best Local Similarity 24.0%; Pred. No. 0.0062;  
 Matches 55; Conservative 19; Mismatches 67; Indels 88; Gaps 12;  
 QY 26 GGGHYIHIRAPSDNLPVLDLIEHPDPIFDPKKDLNETLLRSLLGGHYDPGFMTSPPE 85  
 DB 10 GGDH--HARSTFTVQIP-----FVWDPLDDPAT-----GGCGGFYSP---YSPYS 49  
 QY 86 DRPGG--GGNAGGAEDLAEIDLRLRPS-----GAMP-----SEIKGLE 124  
 DB 50 PYGGGNGAGGAAGGECNLTALRRFLPSNHQDEDEBDGRARGEDVGLGDEFFRYE 109  
 QY 125 FS-EGLAQCKKQRLS-----KKLRKLYQMLMSQTFPCVLYANMDLGRFWRPVYK 174  
 DB 110 FKVRKCARGRSHDWTCEPVAPHPGEKARRDPFRFRFYGTACRPF-----R 154  
 QY 175 VGSCEFSKSGSVPEGM-----VCKPSKSVHLYTLARRCQR 210  
 DB 155 KGACCKGDTCEFAHGVFECWLHPERVRTQACKGOS-----CRRR 194

KW Polyketide; antibiotic Tel-Aviv; cell wall synthesis;  
 KW lipid-disaccharide-pentapeptide; gingivitis.  
 OS Myxococcus xanthus.  
 PN EP1026248-A2.  
 PD 09-AUG-2000.  
 XX 31-JAN-2000; 2000EP-0300747.  
 XX 29-JAN-1999; 99US-0240537.  
 XX (UYRA-) UNIV RAMOT APPLIED RES & IND DEV LTD.  
 PA Rosenberg E, Ron E, Orr E, Patlan Y;  
 PI WPI; 2000-500254/45.  
 DR N-FSDB; AAA59145.  
 XX Novel DNA sequence involved in polyketide antibiotic Tel-Aviv  
 PT production useful for inhibiting cell wall synthesis and in wide range  
 PT of clinical applications such as treating gingivitis  
 XX  
 XX Claim 5; Page 5; 66pp; English.  
 PS The specification describes a DNA sequence which partially encodes  
 CC a functional portion of polypeptide component required for synthesizing  
 CC the polyketide antibiotic Tel-Aviv, postmodification of antibiotic  
 CC Tel-Aviv, or regulation of biosynthesis of antibiotic Tel-Aviv. The  
 CC antibiotic Tel-Aviv is a macrocyclic polyketide synthesised through  
 CC the incorporation of acetate, methionine, and glycine. It inhibits cell  
 CC wall synthesis by interfering with the polymerisation of the  
 CC lipid-disaccharide-pentapeptide. Antibiotic Tel-Aviv genes are useful  
 CC in combinatorial genetics, and for encoding protein components for the  
 CC synthesis, modification and regulation of antibiotic antibiotic Tel-Aviv.  
 CC Antibiotic Tel-Aviv is useful in a wide range of clinical applications  
 CC such as treating gingivitis. Antibiotic Tel-Aviv is also useful for  
 CC generating new biological agents from its secondary metabolites. The  
 CC present sequence represents a protein involved in synthesis of antibiotic  
 CC Tel-Aviv.  
 XX  
 SQ Sequence 2392 AA;  
 Query Match 8.0%; Score 102; DB 21; Length 2392;  
 Best Local Similarity 23.7%; Pred. No. 0.85;  
 Matches 56; Conservative 27; Mismatches 67; Indels 86; Gaps 12;  
 QY 7 LGVTLVAVVVLGRLATPAGGQ-----YHIRAPSDNLPVLDLIEHPDPIFDPKK 59  
 DB 244 LATAIGVFVVMVVRAGL--GQHSFGLRLHL-----DS--VIDGLHAHYPPRVVK 294  
 QY 60 DLNETLLRSLLGGHYDPG-----MATSPEDRPGGGGAGNAGGAEDLAEIDLRLR 109  
 DB 295 D-----LRLSNGEBARAGFTMTFQSLQTLTAPRPERSG-----LPELPLDC 341  
 QY 110 QRPBGAMPSEIKLSESEGLAOCK-----QRLSKLRRKLYQ----- 147  
 DB 342 VHQGAVPLELEVEGAKGLTLFKYDARLYEADYERNAQDLRAADQVADVESPLSA 401  
 QY 148 --WL-----MSQTFCPVLYANMDLGRFWRPVYKVGSCFSKSGSVPEGM 190  
 DB 402 LSWLDEERRTLRLDMNATATPFL---EDLG-----VHELFORCARERTPDAM 445

RESULT 9  
 AAB07665  
 ID AAB07665 standard; Protein; 2392 AA.  
 AC AAB07665;  
 XX  
 XX 07-NOV-2000 (first entry)  
 DT  
 XX Amino acid sequence of a peptide synthetase unit-PKS module.  
 DE  
 XX

RESULT 10  
 AAY53820  
 ID AAY53820 standard; Protein; 413 AA.  
 AC AAY53820;  
 XX  
 XX 13-MAR-2000 (first entry)  
 DT  
 XX



AC AAW22065;  
 XX 07-OCT-1997 (first entry)  
 DT  
 XX  
 DE Chick fringe A (radical) protein.  
 XX  
 KM Fringe A: radical fringe; fate specification; neural tube;  
 KM apical ectodermal ridge; cell proliferation; vasculature;  
 KM atherosclerosis; tumour; wound healing; therapy.  
 XX  
 OS Gallus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 27..28 "a frameshift between codons 27 and 28  
 FT /note= of the fringe A gene results in fringe  
 FT A protein consisting of aa1-27 of one  
 FT open reading frame and aa28 onwards of  
 FT another reading frame"  
 FT Misc-difference 160  
 FT /note= "conserved Cys residue"  
 FT Misc-difference 171  
 FT /note= "conserved Cys residue"  
 FT Misc-difference 189  
 FT /note= "conserved Cys residue"  
 FT Misc-difference 253  
 FT /note= "conserved Cys residue"  
 FT Misc-difference 357  
 FT /note= "conserved Cys residue"  
 FT Misc-difference 366  
 FT /note= "conserved Cys residue"  
 FT  
 XX WO9726276-A2.  
 PN  
 XX 24-JUL-1997.  
 PD  
 XX 15-JAN-1997; 97WO-US00849.  
 XX  
 XX 16-JAN-1996; 96US-0586165.  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 XX  
 XX Laufer EM, Orozco OE, Tabin CJ;  
 PI WPI: 1997-385295/35.  
 DR N-PSDB; AAT75034 and AAT75036.  
 XX  
 XX Chicken fringe A and fringe B protein(s) and DNA - involved in  
 PT differentiation and development of embryonic tissue, used in tumour  
 PT treatment  
 PT  
 XX Claim 9; Fig 1A-B; 64pp; English.  
 PS  
 XX Chicken fringe A (or radical fringe) protein (AAW22065) has the  
 CC ability to affect neural specification of cell type identity,  
 CC proliferative regulation of migratory paths, proliferation of  
 CC neural crest (branchial arches), definition of morphological  
 CC boundaries (including limb), cell shape decisions, regulation of  
 CC apical ectodermal ridge (AER) formation and indirect regulation of  
 CC signalling signals. Fringe A and fringe B (AAW22066) act by  
 CC signalling to or activating members of the Delta-Serrate-Lag-2  
 CC family of transmembrane proteins, which in turn bind to the Notch  
 CC receptor and affect the differentiation and development of  
 CC embryonic tissue. Their amino acid sequences were deduced from  
 CC clones cFR42 (AAT75034) and cFR40 (AAT75035) obtd. from a chick embryo  
 CC library. Fringe proteins can be used in the treatment of tumour  
 CC pathogenesis, disorders of the vasculature system and in wound  
 CC healing, and in assays to identify agents which alter the AER and  
 CC production of angiogenic precursors, the formation of the AER and  
 CC the subdivisions of the neural tube.  
 CC  
 CC Sequence 372 AA;

Query Match 7.4%; Score 93.5; DB 18; Length 372;  
 Best Local Similarity 25.4%; Pred. No. 0.53;  
 Matches 52; Conservative 17; Mismatches 81; Indels 55; Gaps 11;  
 QY 4 CPSLGVTVALVVLGIRATPAGGQHYLIRPAPSNDLPLVLIHPDFIPD--PREKDL 61  
 DB 12 CPLLSTVTAALVLLLLPRGPPAARR---RPPAPR-----SRPSKRARAGSDV 61  
 QY 62 NEYLNSLLGGHYDPGMATSPEDPRPGGGGAAGAEPLAEU---DQLLRPSGAMPS 118  
 DB 62 -----PGDR-GGSGAAGGGRGVAGSPWPSRRVRMGPPGGSAX 98  
 QY 119 EIKGLFESEGLAQKKORLSKLRKLQMWLW-----SQTFCPVLYAMNDLSRFMPRY 173  
 DB 99 E--SLEIKDIFIAVKTTRKHKTRELLPOTMISRAQGTF--IFTWEDRELRL-----L 149  
 QY 174 KVSGCFSKRSCS---VPEGNVCKPS 195  
 DB 150 KAGDHMINTCNSAVHTRQALCKMS 174  
 RESULT 13  
 AAW18658  
 ID AAW18658 standard; Protein; 319 AA.  
 XX  
 XX AAW18658;  
 XX  
 DT 23-JUL-1997 (first entry)  
 DE  
 XX Fragmented human NF-L gene +2 frameshift mutant product.  
 XX  
 KM Frameshift mutation product; GAGA motif; somatic mutation; diagnosis;  
 KM detection; antibody; probe; cancer; neoplasia; neurodegenerative;  
 KM Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;  
 KM Down's syndrome; frontal lobe dementia; progressive supranuclear palsy;  
 KM PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;  
 KM cardiovascular; rheumatoid arthritis; NF-L; neurofilament-low.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 1..319  
 FT /note= "X corresponds to a stop codon in the  
 FT accompanying DNA file, AAT69794"  
 FT  
 XX WO9712992-A2.  
 PN  
 XX 10-APR-1997.  
 PD  
 XX 02-OCT-1996; 96WO-1B01106.  
 XX  
 XX 11-JAN-1996; 96US-0009832.  
 XX 02-OCT-1995; 95GB-0020080.  
 XX  
 XX (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.  
 XX (UYRO-) UNIV ROTTERDAM ERASMUS.  
 XX (UYUT-) UNIV STATE UTRECHT.  
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;  
 DR WPI: 1997-226235/20.  
 DR N-PSDB; AAT69793.  
 XX  
 XX Use of mutant genes having frameshift mutation(s) - for developing  
 PT probe, for the diagnosis, prevention and treatment of associated  
 PT diseases, e.g. cancer or neurodegenerative disease  
 XX  
 PS Claim 22; Fig 7; 123pp; English.  
 XX  
 CC AAW18657 and AAW18658 are +1 and +2 frameshift mutations, respectively,  
 CC of a sequence comprising fragments of the coding sequence of the  
 CC human neurofilament subunit NF-Low (NF-L) gene corresponding to  
 CC nucleotides (nt) 1-420 followed by nt 901-1440 of the wild-type NF-L

CC gene. These regions of the gene contain GAGAG motifs. Frameshift mutants  
CC of the tau, ubiquitin, apolipoprotein E, microtubule-associated protein  
CC 2 (MAP-2), neurofilament subunit L, M and H and amyloid A4 genes are  
CC claimed. All these genes share a common GAGAN motif (N= A, G, C or T),  
CC which is the site of common GA dinucleotide deletion(s) that cause  
CC neurodegenerative disorders. Antigenic peptides used for the production  
CC of antibodies, and small nucleic acid sequences derived from frameshift  
CC mutants are used in the diagnosis, prevention and treatment of cancer  
CC and neurodegenerative diseases, e.g. Parkinson's disease, Alzheimer's  
CC disease, Down's syndrome, frontal lobe dementia (Pick's disease),  
CC progressive supranuclear palsy (PSP), amyotrophic lateral sclerosis,  
CC Huntington's disease, multiple sclerosis, and other degenerative  
CC diseases such as cardiovascular disease and rheumatoid arthritis.

XX Sequence 319 AA;

Query Match 7.3%; Score 92.5; DB 18; Length 319;

Best Local Similarity 23.8%; Pred. No. 0.55;

Matches 66; Conservative 22; Mismatches 110; Indels 79; Gaps 16;

QY 5 PSLGVTLVALVVVGLRATPAGQGHVLRPAPSDNLP-LVDLIEHPDPIFDPKEDLNE 63

DB 27 PLAGVPXSTAGCVTRRRNRGG-----RTAPSRTPQSLPLPSLPCPL-SP----- 72

QY 64 TLLRSLGHHYDGFMA-----TSPED-----RPGGGGAAGABDLAELDQ 106

DB 73 -----GSHRRGAPAAQXVPSATSRTPPTSGATWRRPGCISACAAATAPHAQLTQ 125

QY 107 LLRQR-----PSGAMSEYKGLFSEGLAQKKQRLSKLRKLQWLW---SOTFCPVL 158

DB 126 ATERRCLRPCPCAAAXWTKSL-FXRKCTKRSPNCRRSSTRSPRWWTXPSTP-PPR 183

QY 159 YAWNDLGSFWP-----RYKVGSCFSKR-----SCSVP-----EGMVC---KPS 195

DB 184 SRTSARSTRSWPPTCTRLNGRASRCXPRAPPTPTFCAPPTRCRAVVCSPRPW 243

QY 196 KSVH--LTVLRWRCRRGRCGWIPQIPIISECK 230

DB 244 KSKHAGAXMKWRSSCR-----SWRTSRTPTSALCRC 275

RESULT 14

AAAB10043

ID AAB10043 standard; Protein; 538 AA.

AC AAB10043;

DT 02-NOV-2000 (first entry)

DE MMLV gag protein.

KW Glycoprotein; gag gene; pol gene; GP-1; GP-2; anti-HIV; cytostatic;  
KW gene therapy; treatment; infectious disease; HIV; AIDS; neoplasm;  
KW carcinoma; melanoma.

OS Moloney murine leukemia virus.

PN EP1006196-A2.

PD 07-JUN-2000.

PF 25-NOV-1999; 99EP-0250415.

PR 26-NOV-1998; 98DE-1056463.

PA (PETT-) PETTE INST HEINRICH.

PI Von Laer MD;

XX WPI; 2000-378268/33.

XX New retroviral packing cell useful as pharmaceutical carrier in gene  
PT therapy for treatment of HIV and neoplasms, comprises retroviral genes

PT and glycoproteins -

PS Disclosure; Page 31-32; 69pp; German.

XX This invention describes a novel retroviral packing cell (I), comprising  
CC the retroviral genes gag, pol and glycoproteins GP-1 and GP-2 of the LCMV  
CC coding gene gp, or a part of these. The products of the invention have  
CC anti-HIV and cytostatic activity and can be used for gene therapy. (I) is  
CC useful for in vitro infection of cells, especially hematopoietic stem  
CC cells, for expression of transgenes in cells and as a pharmaceutical  
CC carrier for gene therapy. (II) is therefore useful in the treatment of  
CC infectious diseases (HIV or AIDS), neoplasms (carcinomas), melanomas and  
CC virus (MMLV) gag protein which is described in the method of the  
CC invention.

XX Sequence 538 AA;

Query Match 7.3%; Score 92; DB 21; Length 538;

Best Local Similarity 26.5%; Pred. No. 1.3;

Matches 45; Conservative 12; Mismatches 45; Indels 68; Gaps 9;

QY 21 RATPAGGQHY-----LHIRPAP-----SDN-LPLVDLI-EHPDPIFDPKEDLNETLLRSL 69

DB 122 RSTPPRSSLYPALTPSLGAKPKQVLSGGLIDLLTDEDPYPYRDP----- 169

QY 70 LGGHYDPGFMATSPEDRPGGGGA--AGGAEDLAELDQLLR-----PSGAMPS 118

DB 170 -----PPSDRDGNGGEATPAGEADPPSPMASRLRGRRPPVADSTTSQAFPL 217

QY 119 EIKGLEFSEGLAQKKQRLSKLRKLQWLWSQTFPCVLYAWNDLGSRF 168

DB 218 RAGG-----NGQLQWPFSS--DLYNKNNNPSPF 245

RESULT 15

AAAY17946

ID AAAY17946 standard; Protein; 648 AA.

AC AAAY17946;

DT 03-AUG-1999 (first entry)

DE MoMLV gag gene product.

KW Viral packaging signal; chimeric; type C retrovirus; gag gene; MoMLV;  
KW murine VL30; retroviral vector; biotechnology; pharmaceutical;  
KW gene therapy; Moloney murine leukemia virus.

OS Moloney murine leukemia virus.

PN WO9925862-A2.

XX 27-MAY-1999.

PF 19-NOV-1998; 98WO-US24667.

PR 19-NOV-1997; 97US-0066148.

PA (NATU-) NATURE TECHNOLOGY INC.

PI Guoping X, Hodgson CP, Zink MA;

XX WPI; 1999-347491/29.

DR N-PSDB; AAX77178.

PT New chimeric packaging signal useful in gene therapy

PS Disclosure; Fig 4; 61pp; English.

XX The invention relates to a new chimeric viral packaging signal that  
CC comprises: (i) essential packaging nucleic acid sequence, from a  
CC mammalian type C retrovirus, functionally joined to (ii). (ii) at least

CC one non-essential packaging nucleic acid sequence (lacking the gag gene)  
 CC from murine VJ30. The chimeric viral packaging signals are used to  
 CC produce retroviral vectors for transfer of genetic material in gene or  
 CC cell therapy, or other biotechnological applications. Typical uses are  
 CC in production of pharmaceuticals (e.g. enzymes, antibiotics, antibodies,  
 CC antisense RNA, cytokines etc.) in animals or cell cultures. Elimination  
 CC of the gag gene results in safer vectors (reduced formation of  
 CC replication competent retrovirus as a result of recombination events  
 CC involving the gag gene). Also vectors that contain the chimeric packaging  
 CC signal have high infectious titers (about 1 million transduction units  
 CC (TU)/ml), comparable with that for gag-containing vectors and higher than  
 CC for known gag-free vectors. The present sequence represents the gag gene  
 CC product of the MoMuLV virus.

XX Sequence 648 AA;

Query Match 7.3%; Score 92; DB 20; Length 648;  
 Best Local Similarity 26.5%; Pred. No. 1.6;  
 Matches 45; Conservative 12; Mismatches 45; Indels 68; Gaps 9;

QY 21 RATPAGGQHY-----LHRRAP--SDN-LPLVDLI-EHPDPIFPPEKDLNETLRLSL 69  
 DB 232 RSTPPRSSLYPALTPSLGAKPKQVLSDSGGLIDLTEDPPPYRDPR----- 279  
 QY 70 LGGHVDPGFMATSPBEDRPGGGGA--AGGABDLAELDQLROR-----PSGAMPS 118  
 DB 280 -----PPPSDRDNGGGEATPAGEAPDPSPMASRLRGRRPPVADSTTSQAFL 327  
 QY 119 EIKGLEFSEGLAQKKORLSKLRKLQMWLWSQTFPCPVLYAWNDLGSRP 168  
 DB 328 RAGG-----NGQLQYWPFSSS---DLYNMKNNNSP 355

Search completed: January 7, 2003, 09:23:35  
 Job time : 40 secs